J Neurophysiol 100: 1576–1589, 2008.
First published April 30, 2008; doi:10.1152/jn.01282.2007.

Subthreshold Membrane-Potential Resonances Shape Spike-Train Patterns in the Entorhinal Cortex

T. A. Engel,[1,3] L. Schimansky-Geier,[1,3] A.V.M. Herz,[2,3] S. Schreiber,[2,3,*] and I. Erchova[4,*]

[1]*Department of Physics and* [2]*Department of Biology, Humboldt-Universität zu Berlin;* [3]*Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany; and* [4]*Institute for Adaptive and Neural Computation, University of Edinburgh, Edinburgh, United Kingdom*

Submitted 23 November 2007; accepted in final form 25 April 2008

Engel TA, Schimansky-Geier L, Herz AVM, Schreiber S, Erchova I. Subthreshold membrane-potential resonances shape spike-train patterns in the entorhinal cortex. *J Neurophysiol* 100: 1576–1589, 2008. First published April 30, 2008; doi:10.1152/jn.01282.2007. Many neurons exhibit subthreshold membrane-potential resonances, such that the largest voltage responses occur at preferred stimulation frequencies. Because subthreshold resonances are known to influence the rhythmic activity at the network level, it is vital to understand how they affect spike generation on the single-cell level. We therefore investigated both resonant and nonresonant neurons of rat entorhinal cortex. A minimal resonate-and-fire type model based on measured physiological parameters captures fundamental properties of neuronal firing statistics surprisingly well and helps to shed light on the mechanisms that shape spike patterns: *1*) subthreshold resonance together with a spike-induced reset of subthreshold oscillations leads to spike clustering and *2*) spike-induced dynamics influence the fine structure of interspike interval (ISI) distributions and are responsible for ISI correlations appearing at higher firing rates (≥3 Hz). Both mechanisms are likely to account for the specific discharge characteristics of various cell types.

INTRODUCTION

An ongoing challenge in neuroscience is to relate large-scale network phenomena to the properties of single neurons, including their subthreshold dynamics. It has been shown that subthreshold membrane-potential resonances influence a given brain area's ability to sustain rhythmic activity (Buzsáki and Draguhn 2004; Geisler et al. 2005), which, in turn, underlies various aspects of perception, memory, and behavior (Buzsáki 2005; Lisman 2005). Recent findings for grid cells in layer II of the entorhinal cortex (EC) even suggest that there may be a direct link between the subthreshold dynamics of single neurons and their representation of the sensory environment (Giocomo et al. 2007). Despite the increasing evidence for the interplay of single-cell and network dynamics, it is not fully understood how subthreshold frequency selectivity, as reflected in membrane-potential oscillations and subthreshold resonance, shape spike-discharge patterns in individual neurons. These discharges, however, form the link between single-cell and network activity and are therefore subject to investigation in our study.

Subthreshold resonances are usually examined by injecting an oscillatory current into a neuron, while blocking its synaptic receptors. Intracellular recordings have shown that the voltage response of most neurons, such as EC pyramidal cells, decreases with increasing frequency of the injected current. The cell membrane of these neurons, often referred to as nonresonant, therefore acts as a low-pass filter. In contrast, resonant neurons, such as EC stellate cells (Erchova et al. 2004; Haas and White 2002), have maximum response amplitudes for stimuli delivered at higher frequencies. In the last two decades many examples of resonant neurons have been found, for example, in the inferior olive (Llinás and Yarom 1986), thalamus (Hutcheon et al. 1994), hippocampus (Leung and Yu 1998), and cortex (Gutfreund et al. 1995). These resonant neurons often show subthreshold membrane potential oscillations whose frequencies are close to the resonance frequency.

Subthreshold resonance arises from the activation of delayed membrane currents (as reviewed in Hutcheon and Yarom 2000) and has been shown to influence spiking neural activity by affecting firing rate, precision of spike timing (Desmaisons et al. 1999; Haas and White 2002; Schaefer et al. 2006), and spike clustering (Chen and Shepherd 1997; Desmaisons et al. 1999; Izhikevich et al. 2003; Pedroarena et al. 1999; Wu et al. 2001). Despite these successes, the exact relationship between subthreshold resonance and spike patterns has not yet been established. The current study aims to fill this gap.

Our data underscore the close relation between subthreshold and spiking dynamics, demonstrating that fundamental characteristics of the spike-train statistics in individual neurons can be deduced from their subthreshold resonance properties. Using a resonate-and-fire model with parameters directly constrained by experiments, we quantitatively reproduce the experimentally measured distribution of interspike intervals (ISIs) and their correlations. Importantly, the model provides a straightforward mechanistic explanation for the transition from subthreshold dynamics to spike patterns in both resonant and nonresonant neurons and shows that a spike-induced reset of the phase of subthreshold oscillations accounts for the observed spike clustering in resonant neurons.

METHODS

Physiological data

Experiments, data acquisition, and data analysis were carried out as described in Erchova et al. (2004) and Schreiber et al. (2004a). For the present study, we recorded from 19 neurons (7 stellate, 12 pyramidal) in rat superficial entorhinal cortex. Each neuron was injected with hyper- and depolarizing current pulses lasting for 500 ms, 10 s, 20 s, and 30 s to characterize the cell type and to obtain ISI distributions.

* These authors contributed equally to this work.

Address for reprint requests and other correspondence: I. Erchova, Institute for Adaptive and Neural Computation, University of Edinburgh, Informatics Forum, Crichton St., FH1 9AR (E-mail: ierchova@inf.ed.ac.uk).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

To study resonance properties, cells were injected with an oscillatory "ZAP" current of 30-s duration

$$I(t) = I_0 + I_1 \sin[2\pi f(t)t] \quad (1)$$

whose oscillation frequency increased in time, $f(t) = f_m t/2T$ (Gimbarzevsky et al. 1984; Puil et al. 1986). Here, I_1 denotes the stimulus amplitude, I_0 is the injected constant current, T is the stimulus duration, and $f_m = 20$ Hz is the maximum stimulus frequency.

The true membrane potential of all recorded cells was verified at the end of each recording session. The cell-voltage offset was measured after the electrode was gently pulled out of the cell and cleaned from the attached membrane. All recordings were then corrected by the value of this offset.

Resonate-and-fire model

SUBTHRESHOLD REGIME. To describe subthreshold dynamics we use an extension of an electrical circuit model that has previously been successfully used to capture the frequency-dependent subthreshold dynamics in stellate and pyramidal cells of the EC (Erchova et al. 2004; Schreiber et al. 2004a). In the circuit interpretation, the original model consists of two parallel branches (Fig. 1*A*). The first branch is characterized by a conductance g in parallel with a capacitance C and mimics the integrative properties of a leaky integrator. The second branch consists of a resistance R_L in series with an inductance L and captures the response properties of delayed rectifying currents.

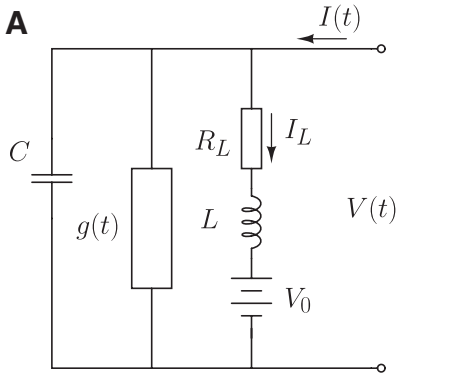
In the present study, we model the dynamics driven by cell-intrinsic noise. Because synaptic transmission was blocked, the dominant intrinsic noise source was channel noise (Manwani and Koch 1999; White et al. 2000). Accordingly, the conductance becomes a time-dependent quantity $g(t)$ whose dynamics we describe by an additional stochastic differential equation (Fig. 1*A*). This equation captures the dynamics of a population of ion channels that open and close independently and with constant rates. Although the stochastic conductance $g(t)$ creates multiplicative noise, the full three-dimensional model can be reduced to a two-dimensional dynamical system with an additive noise term $\xi(t)$. This simplification (derived in Verechtchaguina et al. 2007) is possible because the correlation time of conductance fluctuations ($\ll$1 ms; see White et al. 1998) is much smaller than the dominant timescale of the subthreshold voltage dynamics (10–100 ms). The membrane potential $V(t)$ thus evolves according to

$$C\frac{d^2V(t)}{dt^2} + \gamma\frac{dV(t)}{dt} + \delta V(t) = \delta[R_0 I(t) + V_r] + \frac{dI(t)}{dt} + \sqrt{2D}\xi(t) \quad (2)$$

Here, $I(t)$ represents the time-dependent input current and $\xi(t)$ is white Gaussian noise ($\langle\xi(t')\rangle_{t'} = 0$, $\langle\xi(t + t')\xi(t')\rangle_{t'} = \delta(t)$) where $\delta(t)$ is Dirac's function.

In the model, each cell is characterized by its membrane capacity C, resting membrane potential V_r, membrane input resistance R_0, and intrinsic noise intensity D. Resonance properties are described by two additional parameters, γ and δ. In terms of damped harmonic oscillators, γ/C represents the effective damping coefficient and $(1/2\pi)\sqrt{\delta/C}$ is the effective eigenfrequency. The relationship between the parameters of the circuit model and the parameters V_r, R_0, γ, and δ used here is summarized in Fig. 1*A*.

In the noise-free model ($D = 0$), stimulation with a short current pulse results in a damped subthreshold oscillation when $\gamma < 2\sqrt{\delta C}$ (underdamped regime). The frequency of the resulting oscillation is $f_{osc} = \Omega/2\pi = (1/2\pi)\sqrt{\delta C - \gamma^2/4}/C$ and the amplitude decays exponentially with a time constant $t_{rel} = 2C/\gamma$. For $\gamma > 2\sqrt{\delta C}$ (overdamped regime), the voltage V relaxes monotonically to the resting potential. Driven by intrinsic noise ($D > 0$), the model generates sustained subthreshold oscillations with fluctuating amplitude and phase in the underdamped regime ($\gamma < 2\sqrt{\delta C}$), whereas in the overdamped regime ($\gamma > 2\sqrt{\delta C}$) the subthreshold voltage fluctuations do not exhibit rhythmic components.



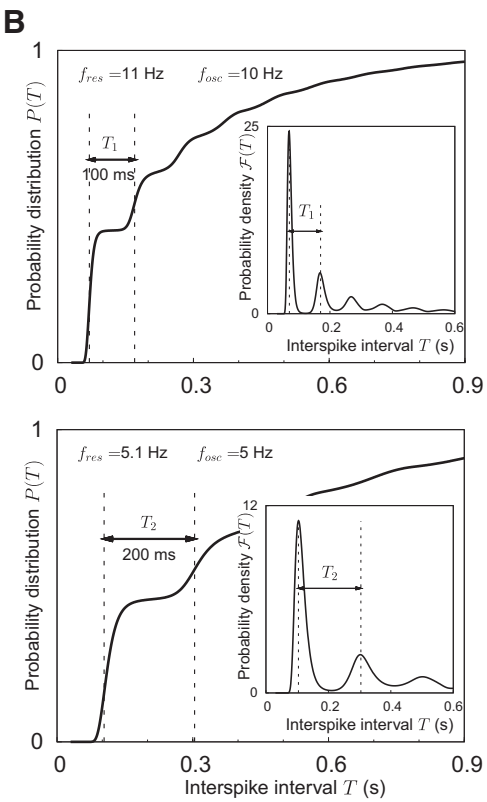


FIG. 1. Schematic representation of the subthreshold model part and interspike interval (ISI) distributions obtained with the renewal model. *A*: the subthreshold part of the model extends the electrical circuit analogy of Erchova et al. (2004) and Schreiber et al. (2004a). Specifically, the previously constant conductance of the middle parallel branch is now replaced by a time-dependent stochastic conductance to capture cell-intrinsic noise. As stated, the parameters γ, δ, R_0, and V_r can be expressed in terms of the circuit parameters. The parameter Q denotes the noise strength of conductance fluctuations: $Q \sim Np(1 - p)/\tau$, where N is the number of channels, p is the opening probability, and τ is the correlation time of conductance noise. For details, see Verechtchaguina et al. (2007). *B*: two cumulative ISI distributions obtained with the renewal model (in the Stratonovich approximation). The analytical curves are obtained for two parameter sets that differ in the subthreshold resonance frequency f_{res} and, consequently, also in the frequency of subthreshold membrane potential oscillations f_{osc} (*top*: f_{osc} = 10 Hz; *bottom*: f_{osc} = 5 Hz); see *Eq. 4*. The distance between the 1st and 2nd peaks of the ISI densities (*insets*) is equal to the period of subthreshold oscillations and thus also approximately equal to the inverse of the subthreshold resonance frequency.

To relate the resonance properties of the model to experimental data, the theoretical frequency-dependent impedance function $|Z|(f)$ was calculated for the noise-free case (i.e., *Eq. 2* with $D = 0$). This simplification mimics the experimental situation where impedances were calculated from voltage responses averaged over at least five trials, and results in

$$|Z_{\text{theory}}|(f) = \sqrt{\frac{4\pi^2 f^2 + R_0^2\delta^2}{[\delta - 4\pi^2 C f^2]^2 + 4\pi^2\gamma^2 f^2}} \tag{3}$$

In the underdamped regime the impedance function has a maximum at the frequency f_{res}

$$f_{\text{res}} = \frac{1}{2\pi}\sqrt{\frac{\delta}{C}}\sqrt{\sqrt{(1 + C\delta R_0^2)^2 - \gamma^2 R_0^2} - C\delta R_0^2} \tag{4}$$

For small γ, the resonance frequency f_{res} is similar to the effective eigenfrequency and the frequency f_{osc} of the deterministic subthreshold oscillation, with f_{res} slightly exceeding f_{osc}. Recordings from stellate cells of the entorhinal cortex (Erchova et al. 2004) confirm this prediction of our overall model framework. In the overdamped regime the impedance function $|Z_{\text{theory}}|(f)$ decays monotonically with increasing frequency (i.e., the cell is nonresonant). Nevertheless, it may have a nonzero eigenfrequency (Erchova et al. 2004).

In the model, injecting a prolonged constant current $I(t) = I_0$ shifts the stationary voltage to a new value $V_r + I_0R_0$. Since we were interested only in the stationary responses we introduced an auxiliary variable $x(t) = V(t) - V_r - I_0R_0$, which describes fluctuations around a zero-mean voltage

$$C\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} + \gamma\frac{\mathrm{d}x(t)}{\mathrm{d}t} + \delta x(t) = \sqrt{2D}\xi(t) \tag{5}$$

Already in its noise-free form, this simple phenomenological model reproduces both resonant and nonresonant subthreshold dynamics (Erchova et al. 2004; Izhikevich 2001; Mauro et al. 1970; Richardson et al. 2003; Verechtchaguina et al. 2006b) and can be related to Hodgkin–Huxley-type conductance-based models via linearization (Mauro et al. 1970; Villacorta and Panetsos 2005). Because the model parameters depend in general on the level of membrane depolarization, experimental data for each value of the input current I_0 were treated separately, similar to the approach taken by Erchova et al. (2004). The range of parameters was then estimated for each I_0 value (as described in the following text).

SPIKING REGIME. *Renewal model.* For a minimal description of spike generation, the model needs to be equipped with a firing threshold and a voltage reset. Slightly extending this approach to incorporate refractoriness, we introduced three parameters describing spike generation: voltage threshold x_b, voltage reset x_0, and reset time τ_r. Whenever $x(t)$ reaches x_b from below, a spike is generated, and $x(t)$ is reset to x_0, $\mathrm{d}x(t)/\mathrm{d}t$ to zero. After the reset period τ_r, the time evolution of x again follows *Eq. 5*. For $D > 0$, this dynamics generates a sequence of independent spikes and is referred to as the *renewal* resonate-and-fire model. Two examples of ISI distributions generated with this model are shown in Fig. 1*B*. Analyzed in detail in RESULTS, one of the model predictions is that the distance between the first and the second peaks in the ISI density is equal to the period of subthreshold oscillations. The inverse of this peak-to-peak distance is therefore approximately equal, but slightly smaller than the subthreshold resonance frequency (see explanation following *Eq. 4*). The change in peak-to-peak distance with the subthreshold resonance and oscillation frequencies is illustrated in Fig. 1*B*.

Model extension: nonrenewal process. To cover ISI correlations, deterministic spike generation was replaced by a stochastic rule. This means that a spike is not necessarily generated when the membrane potential crosses the voltage threshold. Rather, spike generation fails with a probability $p_s(t)$

$$p_s(t) = \min\left[1, \sum_i p_0\frac{1}{2}\{1 - \tanh[\beta(t - t_i - \tau_c)]\}\right] \tag{6}$$

Here, t_i represents the spike times and the summation is taken over all spikes that have been generated before time t. If the expression under the sum exceeds 1, then $p_s(t) = 1$ and no spike can be generated.

The probability $p_s(t)$ is the key new dynamic variable. Initially, $p_s(t)$ is set to zero. When triggered by a single spike, $p_s(t)$ instantaneously increases by an amount p_0, remains roughly at this level during the correlation time τ_c, and then decays back to zero on a timescale $2/\beta$. The probability $p_s(t)$ is evaluated each time the voltage variable $x(t)$ crosses the threshold x_b from below; with probability $1 - p_s(t)$ a spike is generated. If no spike is elicited, $x(t)$ evolves without reset according to *Eq. 5* until it again hits the threshold from below. If several spikes are generated within a time window smaller than the correlation time, the probability $p_s(t)$ exhibits complex dynamics as illustrated in Fig. 6. As the example suggests, the stochastic threshold mechanism limits the number of spikes within a cluster and causes "silent overshoots" (allowing the membrane potential to briefly exceed the threshold value without producing a spike). The resulting *nonrenewal* resonate-and-fire model generates spike trains whose ISIs are negatively correlated over short timescales. The prediction that the inverse distance between the first and second peaks in the ISI density is approximately equal to the subthreshold resonance frequency also holds for the nonrenewal model.

PARAMETER ESTIMATION. *Subthreshold dynamics.* The model parameters in *Eq. 5* were estimated from impedance measurements and autocorrelation analysis. The frequency-resolved impedance was determined from the voltage response $V(t)$ to the injected ZAP current $I(t)$ (*Eq. 1*) as $Z(f) = \text{FFT}[V(t)]/\text{FFT}[I(t)]$, where FFT is the fast Fourier transform algorithm. Using the analytical expression for $|Z_{\text{theory}}|(\text{f})$ (*Eq. 3*), the values of the parameters C, δ, γ, and R_0 were obtained from least-square fits of $|Z_{\text{theory}}|(f)$ to the experimental curve $Z(f)$. For two examples, see Fig. 2, *A* and *B*. For larger ZAP amplitudes and at more depolarized levels of membrane potential occasional spikes were observed in the ZAP response. For impedance estimation, spikes were removed by cutting voltages at −55 mV, which did not change the impedance estimates. The estimates of resonance frequency presented in Fig. 8 are obtained as averages of resonance frequencies across different levels of membrane depolarization (based on model fits to the data). Given that resonance frequencies in stellate cells are relatively constant with depolarization (Erchova et al. 2004), this procedure is justified.

The intrinsic noise intensity D was estimated by comparing the experimental and theoretical autocorrelation function $r_x(t) = \langle x(t + t')x(t')\rangle_{t'}$. The experimental autocorrelation function was calculated from voltage responses to hyper- and depolarizing current steps, for time intervals that did not contain spikes. A Butterworth low-pass filter (cutoff frequency 40 Hz) was applied to the data before the parameter estimation. The autocorrelation function of the model can be obtained analytically. In the underdamped regime ($\gamma < 2\sqrt{\delta C}$) it reads

$$r_x(t) = \frac{D}{\gamma\delta}e^{-\frac{\gamma}{2C}|t|}\left[\frac{\gamma}{2\Omega C}\sin(\Omega|t|) + \cos(\Omega t)\right] \tag{7}$$

where $\Omega = \sqrt{\delta C - \gamma^2/4}/C$. Accordingly, $r_x(t)$ oscillates with the period $T_p = 2\pi/\Omega$ and its amplitude decays with the relaxation time $t_{rel} = 2C/\gamma$ (see Fig. 2*C*).

Using *Eq. 7*, the values of $r_x(t)$ at $t = 0$ and at the first minimum t_1, as well as the value of C (from the impedance estimate), we calculated parameter values for stellate cells via the following relations

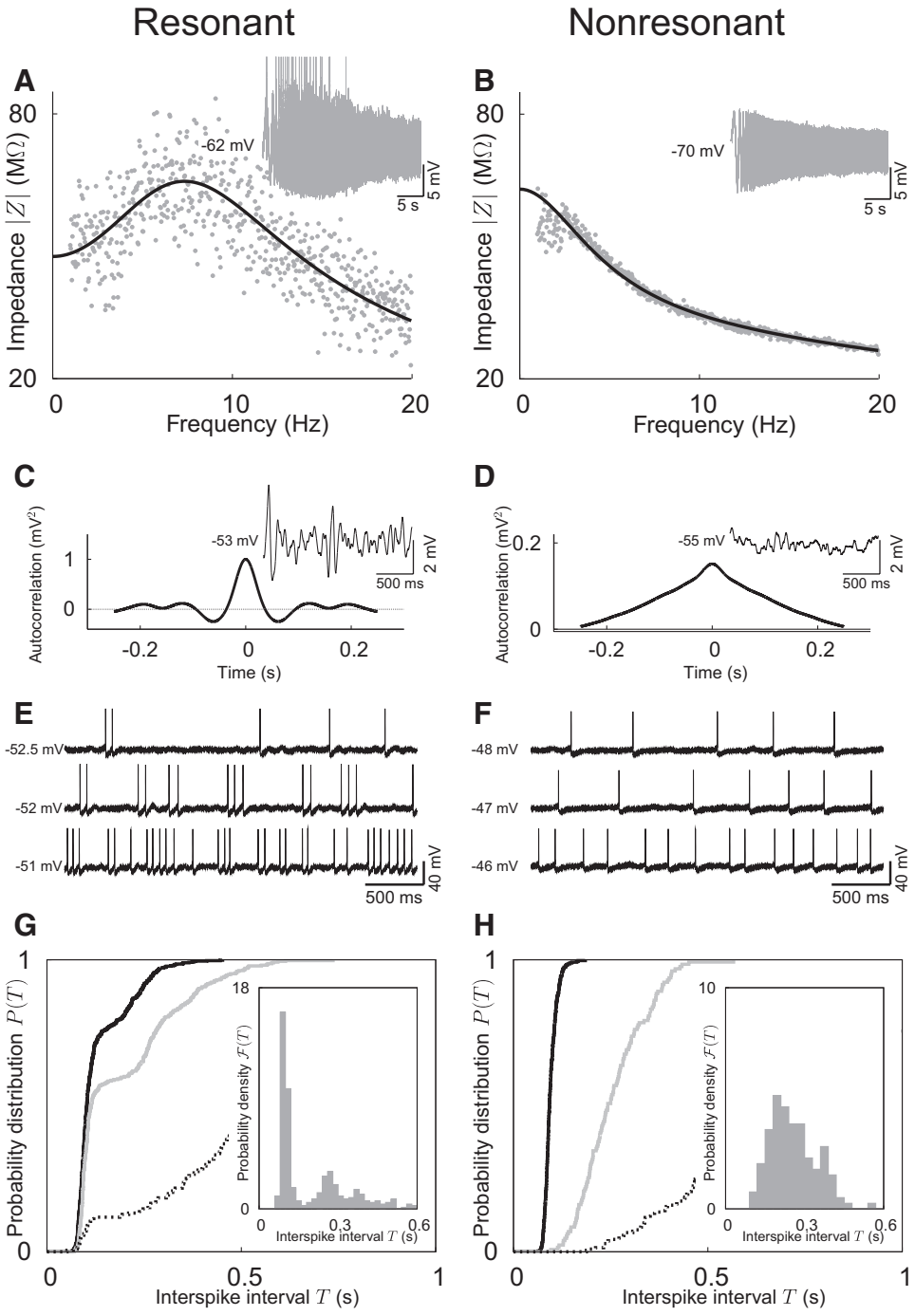


FIG. 2. Subthreshold dynamics and firing patterns of stellate and pyramidal cells in the rat entorhinal cortex (*left* and *right columns*, respectively). *A* and *B*: impedance-amplitude profiles (main panels) as calculated from voltage responses (*insets*) to time-varying oscillatory (ZAP) currents with amplitude $I_1 = 100$ pA. The stellate cell has a clear resonance maximum at about 7 Hz; the pyramidal cell is nonresonant (gray: experiment; black: model fit). *C* and *D*: autocorrelation functions (main panels) of subthreshold voltage responses (*insets*) to constant current injections. Rhythmic components of the membrane-potential oscillations cause the trough and side peaks in the autocorrelation function of the stellate cell ($I_0 = 250$ pA). In the pyramidal cell ($I_0 = 300$ pA), on the other hand, the autocorrelation function decays monotonically, as expected for nonrhythmic fluctuations. *E* and *F*: spike-train responses to constant depolarizing currents in the stellate ($I_0 = 300$, 400, and 650 pA) and pyramidal cell ($I_0 = 600$, 700, and 900 pA). The stellate, but not the pyramidal cell, generates clustered action potentials. *G* and *H*: ISI statistics. Cumulative ISI distributions $P(T)$ are shown for 3 different activity levels (with firing rates of 1.5, 5.1, and 7.5 Hz for the stellate cell; 1.3, 3.9, and 10.4 Hz for the pyramidal cell, respectively). Probability densities $\mathcal{F}(T)$ are depicted in the *insets* for the intermediate firing rates. In the stellate cell, $P(T)$ exhibits a plateau structure. Consequently, $\mathcal{F}(T)$ is multimodal. For all firing rates the sharp rise in $P(T)$ occurs at about the same T. This interval corresponds to the dominant ISI within a spike cluster. For the pyramidal cell, the ISI probability density is unimodal and, correspondingly, $P(T)$ has a single inflection point.

$$\text{(a)}\quad \Omega = \frac{\pi}{t_1} \qquad \text{(b)}\quad \gamma = \frac{2C}{t_1}\ln\left[-\frac{r_x(0)}{r_x(t_1)}\right]$$
$$\text{(c)}\quad \delta = C\Omega^2 + \gamma^2/(4C) \qquad \text{(d)}\quad D = r_x(0)\gamma\delta \qquad (8)$$

In the overdamped regime (when $\gamma > 2\sqrt{\delta C}$), the autocorrelation function $r_x(t)$ is given by a variant of *Eq. 7*, where the trigonometric functions are replaced by their hyperbolic counterparts. Consequently, $r_x(t)$ decays monotonically (Fig. 2*D*).

All parameters were estimated separately for each steady-state depolarization. Note that parameters do not vary independently, but obey an additional constraint, $\sigma_x^2 = D/(\gamma\delta)$, imposed by the model. For resonant cells γ and δ were estimated in two ways, using either impedance or autocorrelation functions. The values obtained by both methods were almost identical.

Spike dynamics. The parameters of the renewal model describing spike generations—i.e., voltage threshold V_b, reset value V_0, and reset time τ_r—were directly estimated from the experimental data. To determine the value of the spike threshold V_b, we first estimated the SD σ_{VD} of the voltage derivative in the absence of spikes. Second, we defined a derivative threshold for spike initiation as $3\sigma_{VD}$. Third, we defined the local voltage threshold V_b as that point where the voltage derivative crosses the derivative threshold. The procedure was performed for each individual spike. The local reset value V_0 was set to the minimal voltage within 25 ms after the spike, whereas the local reset duration τ_r was defined as the time between threshold crossing and reaching of the reset value. All parameters were first evaluated for each spike and then averaged over all spikes generated at a given value of membrane depolarization (not less than 40 spikes). The

steady-state voltage was estimated by the mean membrane potential in time intervals without spikes. The model threshold x_b and voltage reset value x_0 were obtained by subtraction of the steady-state voltage from V_b and V_0, respectively.

Spike-train statistics

ISI DISTRIBUTIONS. We denoted successive spike times by $t_1, t_2, \ldots, t_i, \ldots, t_{N+1}$, where i is the position of a spike in the measured spike train. The ith ISI was defined as $X_i = t_{i+1} - t_i$.

A rough estimate for the ISI probability density $\mathcal{F}(T)$ can be obtained from binned experimental data. The inverse distance between the first and second peaks in the ISI densities in Fig. 8 is based on distances averaged across ISI densities obtained for two to three different values of injected constant current. To adequately localize the second peak, only ISI distributions based on firing rates of ≥4 Hz were considered. This criterion was satisfied by six of seven stellate cells. The ISI densities were sampled with a bin width of 15 ms and peaks were identified at bins with the largest amplitude.

Clustering in stellate cells was reflected in multimodal ISI distributions, as shown in Fig. 2*G*. For this study, we therefore defined clustered spike patterns as those patterns where the corresponding ISI distribution showed a significant minimum between its first two peaks. Specifically, we required a confidence level of ≥95%, based on a likelihood-ratio test. Spikes separated by a distance smaller than the distance at this minimum were considered as belonging to a cluster; typical distances between spike clusters corresponded to the location of the second and potentially third peaks of the ISI density. Note that our criterion of multimodal ISI distributions is stronger than what is theoretically needed to discriminate between clustered and unclustered spike trains; however, all stellate cells with sufficient statistics satisfied this criterion. No pyramidal cells showed spike clustering, as additionally confirmed by visual inspection.

Due to the limited sample size (150–600 spikes per spike train) the exact shape of estimated ISI densities depends strongly on the bin size. For model fitting, we therefore used the cumulative ISI distribution $P(T) = \int_0^T \mathcal{F}(t)\mathrm{d}t$. Unlike $\mathcal{F}(T)$, $P(T)$ can be reliably estimated from limited data as

$$P_X(T) = \frac{1}{N}\sum_{i=1}^{N} \Theta(T - X_i) \tag{9}$$

where $\Theta(T)$ is the Heaviside step function [$\Theta(T) = 0$ if $T < 0$ and $\Theta(T) = 1$ if $T > 0$].

ISI CORRELATIONS. Because the same ISI distribution could potentially be generated by a renewal or a nonrenewal process, we used the following measures to test whether ISIs were independent:

1) The joint probability density $\mathcal{F}(T_1, T_2)$ of two successive ISIs (i.e., the density of the ISI return map). $\mathcal{F}(T_1, T_2)\mathrm{d}T_1\mathrm{d}T_2$ gives the probability that successive ISIs satisfy $X_i \in (T_1, T_1 + \mathrm{d}T_1)$ and $X_{i+1} \in (T_2, T_2 + \mathrm{d}T_2)$. The joint densities $\mathcal{F}(T_1, T_2)$ were estimated by histograms and normalized to unity.

2) Correlation maps $C(T_1, T_2)$, characterizing the deviation of the experimental ISI sequence from the renewal ISI sequence (with independent ISIs). Correlation maps $C(T_1, T_2)$ were obtained from measured joint ISI densities by subtracting the joint ISI density expected for the corresponding renewal spike train

$$C(T_1, T_2) = \mathcal{F}(T_1, T_2) - \mathcal{F}(T_1)\,\mathcal{F}(T_2) \tag{10}$$

Positive values of $C(T_1, T_2)$ indicate that this ISI pair occurs more frequently than under the renewal assumption; negative values indicate pairs that are less frequent.

3) Serial correlation coefficients (SCCs) ρ_k describe correlations between arbitrarily distant ISIs

$$\rho_k = \frac{\langle X_i X_{i+k}\rangle - \langle X_i\rangle\langle X_{i+k}\rangle}{\langle X_i^2\rangle - \langle X_i\rangle^2} \tag{11}$$

Here, averages are taken over all ISIs in a spike train and the index k is called a lag. For a renewal spike train $\rho_k = 0$ for all $k > 0$. To estimate the typical correlation time, we plot the SCCs as a function of time, by linearly scaling the lag axis with the mean ISI.

MODEL OPTIMIZATION. To find a set of model parameters reproducing the experimental ISI distribution, we estimated the permitted parameter range from the experimental data (as described earlier). The required parameters were found through a simulated annealing (SA) technique (Press et al. 1999) described in the supplemental material.[1]

The membrane capacity C was held constant ($C = 2.5 \times 10^{-4}$ μF for the stellate cell and $C = 1.7 \times 10^{-4}$ μF for the pyramidal cell) because it typically did not depend on the membrane potential; γ, δ, and D were optimized within the permitted parameter range (see Fig. 3 and Table 1). Likewise, the threshold x_b and reset x_0 varied within the measurement range. Since V_r depended sensitively on the threshold definition, we tested an extended range of thresholds to compensate for possible biases. To capture spike-induced effects potentially lasting beyond the time of maximum hyperpolarization, the reset time τ_r was allowed to vary up to the size of the minimal ISI observed in each specific spike train. The parameters for the nonrenewal threshold mechanism—τ_c, β, and p_0—were allowed to vary freely within the range indicated in Table 1. Note that optimal parameter values for the nonrenewal model differed from the renewal model (Table 1). The nonrenewal model admits "silent" threshold crossings without spike generation. To obtain similar firing rates, the total number of threshold crossings was therefore larger for the nonrenewal model.

To quantify the discrepancy between the measured ISI distribution $P_X(T)$ and the theoretical distribution $P_0(T)$, we used the Kolmogorov–Smirnov (KS) test (Ledermann 1984). The test hypothesis was that the theoretical distribution $P_0(T)$ is the distribution underlying the measured ISIs. The test statistics $\mathcal{D}$ in the KS test represents the largest difference between two distributions

$$\mathcal{D} = \sup_{0<T<\infty} |P_X(T) - P_0(T)| \tag{12}$$

For a significance level $\alpha = 5\%$ and a given sample size N, the critical value $\mathcal{D}_K$ is calculated as $\mathcal{D}_K = 1.36/\sqrt{N}$. If $\mathcal{D}$ exceeds the critical value $\mathcal{D}_K$, the hypothesis has to be rejected. To graphically illustrate the KS test, a tube with radius $\mathcal{D}_K$ was drawn around the curve corresponding to the experimental ISI distribution $P_X(T)$. If the theoretical ISI distribution $P_0(T)$ laid completely inside this tube, it was regarded as a potential distribution underlying the measured ISIs.

For the renewal model, the ISI distribution $P_0(T)$ was obtained analytically using the Stratonovich approximation (Verechtchaguina et al. 2006a,b) described in the supplemental material. Because the model parameters for both resonant and nonresonant cells belong to a regime of moderate damping and moderate noise intensity, the Stratonovich approximation was applicable and accurate. For the nonrenewal model, $P_0(T)$ was estimated on the basis of stochastic simulations consisting of ≥3,000 spikes.

The SA procedure is described in detail in the supplemental material. For the renewal model the procedure minimizes the KS statistics $\mathcal{D}$. For the nonrenewal model the procedure simultaneously minimizes the KS statistics $\mathcal{D}$ and the discrepancy between experimentally measured and theoretically predicted values of the first serial correlation coefficient ρ_1.

Note that SA belongs to a class of heuristic algorithms that do not necessary converge to the global optimal solution. Thus the method allows us to find a model generating a spike pattern similar to that observed experimentally, but does not guarantee that this model is unique.

[1] The online version of this article contains supplemental data.

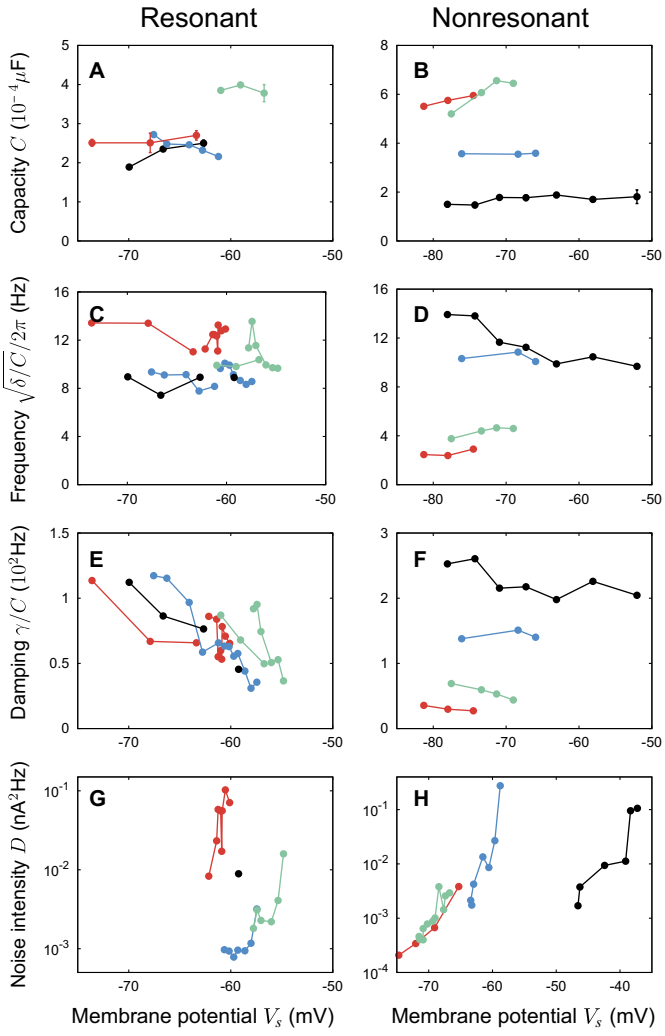


FIG. 3. Resonate-and-fire model. Voltage dependence of subthreshold model parameters, as obtained from physiological data. The 4 rows represent the four parameters (C, $\sqrt{\delta/C}/2\pi$, γ/C, D), shown for 4 stellate and 4 pyramidal cells (*left* and *right columns*, respectively). Different colors correspond to different neurons; the 2 cells presented in Fig. 1 are marked in black. *A* and *B*: capacity C, derived from fitting the experimental impedance function with the theoretical curve obtained from the resonate-and-fire model. *C* and *D*: eigenfrequency $\sqrt{\delta/C}/2\pi$, obtained from fitting the experimental impedance and autocorrelation functions with the model. *E* and *F*: damping coefficient γ/C, calculated from the experimental impedance and autocorrelation functions. *G* and *H*: noise intensity D, obtained from the experimental autocorrelation function. As shown by these data, the noise intensity D strongly depends on the level of membrane depolarization in both cell classes. In addition, in stellate cells the damping coefficient γ/C decreases rapidly with increasing membrane depolarization.

RESULTS

This study aims to elucidate the relation between subthreshold and spiking dynamics of single neurons. To this end, we performed intracellular recordings of stellate and pyramidal cells, two major cell classes in the entorhinal cortex. Constant and time-varying ZAP-current injections (see METHODS) were used to investigate spontaneous subthreshold membrane-potential oscillations and resonances, respectively. Furthermore, constant depolarizing currents were applied to move the mean membrane potential near threshold and thus elicit spike trains driven by cell-intrinsic noise. The data were interpreted within the framework of resonate-and-fire model neurons.

Subthreshold dynamics

Stellate cells exhibited a pronounced membrane-potential resonance in the theta range (5–15 Hz), whereas pyramidal cells showed mostly none, or at most a very small resonance at lower frequencies, between 2 and 4 Hz (Fig. 2, *A* and *B*). In all resonant cells, small constant depolarizing currents caused voltage oscillations with a rhythmic component close to the resonance frequency (Fig. 2*C*). In contrast, the voltage fluctuations of nonresonant cells did not exhibit any rhythmicity (Fig. 2*D*).

Firing patterns

For larger depolarizing currents all investigated neurons produced action potentials. The discharge patterns, however, differed significantly between both cell types. In particular, clustered spikes were observed only for stellate cells (Fig. 2, *E* and *F*). With increasing firing rate, an increasing number of spikes belonged to a cluster (increased plateau level in the ISI distribution in Fig. 2*G*) and average intervals between clusters decreased (shortened tail of the ISI density). Remarkably, the dominant intracluster ISI of stellate cells did not vary with firing rate. In contrast to stellate cells, the ISI density in pyramidal cells was monomodal and the cumulative ISI distribution lacked any plateaus. Higher firing rates simply caused a shift of the distribution to smaller ISI values (Fig. 2*H*).

To summarize, spike patterns in stellate cells exhibited two different timescales reflecting inter- and intracluster ISIs. Increasing the firing rate decreased the intercluster ISIs but did not affect intracluster ISIs. Pyramidal cells, on the other hand, had only a single timescale that directly corresponded to their firing rate.

Resonate-and-fire model

As the first step of our computational analysis, we reproduced the experimental ISI distributions with a minimal resonate-and-fire model (see METHODS). Its subthreshold dynamics are given by a linear differential equation that covers both resonant and nonresonant membrane impedances. As shown by previous studies this approach captures the different subthreshold responses of stellate and pyramidal cells surprisingly well, despite its minimal complexity. In addition, the model makes quantitative predictions that have been tested experimentally (Erchova et al. 2004; Schreiber et al. 2004a).

Here, we extend this successful approach to the spiking regime and investigate whether the model also provides a mechanistic understanding of the spike-train statistics of different types of entorhinal cortex cells. To this end, it is important to recapitulate that the parameters of this phenomenological description are allowed to change with the membrane potential so as to account for voltage-dependent cell properties. For each level of membrane depolarization, one set of parameters was therefore estimated (Fig. 3). In both stellate and pyramidal cells, the effective capacity C and eigenfrequency $\sqrt{\delta/C}/2\pi$ depended only weakly on the membrane potential (Fig. 3, *A–D*). For the purpose of this study, C was thus considered constant. In resonant cells the damping constant γ/C decreased with depolarization (Fig. 3*E*), whereas it was approximately constant in nonresonant cells (Fig. 3*F*). This result agrees with the observation that the amplitude of sub-

TABLE 1. *Parameters of resonate-and-fire models that reproduce the experimental ISI distributions of the stellate cell and pyramidal neuron shown in Figs. 1, 4, and 7*

	$\sqrt{\delta/C}/2\pi$, Hz	γ/C, 10^2 Hz	D, nA^2 Hz	τ_r, ms	x_0, mV	x_b, mV	τ_c, ms	β, Hz	p_0
				A. Sample stellate cell					
Range	5–7.5	0.02–0.6	0.005–0.05	0–51	−12 to −7	1–7	100–500	1–1,000	0–1
Renewal (i)	6.2	0.09	0.027	35	−7.4	5.8	—	—	—
Renewal (ii)	7.1	0.45	0.015	30	−8.4	1.9	—	—	—
Nonrenewal	6.8	0.28	0.009	44	−8.2	1.0	322	736	0.22
				B. Sample pyramidal cell					
Range	7–14	1.8–3.5	0.05–0.5	0–97	−12 to −7	0.02–6.0	100–500	1–1,000	0–1
Renewal	11.1	3.24	0.219	74	−10.9	0.3	—	—	—
Nonrenewal	13.9	2.09	0.124	87	−11.4	0.5	341	156	0.41

Spiking parameters were estimated at firing rates between 1 and 10 Hz.

threshold resonances increases with depolarization (Erchova et al. 2004). Finally, in both cell classes the intrinsic noise intensity D increased strongly with membrane depolarization (note the logarithmic scale in Fig. 3, *G* and *H*), in agreement with findings of Diba et al. (2004) in cultured hippocampal neurons.

To include spike generation in the model without increasing its complexity more than absolutely necessary, the subthreshold dynamics were combined with a constant voltage threshold x_b. When the membrane potential reaches x_b, an action potential is elicited, the membrane potential is reset to a fixed value x_0, and its derivative $\mathrm{d}x(t)/\mathrm{d}t$ is reset to zero. To account for refractoriness, the subthreshold dynamics begins again after a fixed reset time τ_r. These three additional parameters were also estimated from the experimental traces (see Table 1).

ISI distributions are captured by the resonate-and-fire model

The subthreshold ("resonate") and spiking ("fire") parameters of the resonate-and-fire model are obtained at different levels of depolarization. Since the subthreshold parameter values depend on the membrane potential, they cannot be directly inserted into the model. We rather used the ranges of parameters measured across different membrane potentials to constrain the parameter space and searched within this space for the optimal parameter set by simulated annealing (see supplemental material). The criterion of optimality was defined as the lowest KS distance (see METHODS) between experimental and model ISI distributions. The latter were obtained analytically as the distributions of first passage times using the Stratonovich approximation (as described in the supplemental material). Optimized model ISI distributions are consistent with the experimental data for both cell types (Fig. 4).

Structurally different ISI distributions were obtained with different parameter sets, as shown in Fig. 4*A*. Here, all model parameters were either varied independently (solid line) or constrained by a fixed upper bound for the variance of the subthreshold voltage fluctuations σ_x^2 (dashed line). Although both optimizations captured the main ISI peak, the first parameter set reproduced the multimodal probability density better than the second. However, the better fit came at the cost of subthreshold voltage fluctuations much larger than observed in experiments (first model set: $\sigma_x = 5.6$ mV; second model set: $\sigma_x = 1.6$ mV; experiments $\sigma_x \approx 2.5$ mV).

These results can be understood from considering the model features that shape the ISI density. In both experiment and model, spikes are more likely to occur at the maxima of subthreshold oscillations, where the distance to threshold is smallest (see also Alonso and Klink 1993; Klink and Alonso 1993). Within the resonate-and-fire model, the phase of subthreshold oscillations is reset to a fixed value after each spike. Consequently, the initial subthreshold oscillation following a spike is similar for all spikes (despite the presence of noise) and the membrane potential reaches its first maximum roughly at the same time lag T, for small to moderate noise levels. T corresponds to the intracluster ISI and is equal to the sum of the reset period τ_r and the time between the end of the reset period and the first maximum of the following subthreshold oscillation (occurring at approximately half an oscillation period). The second peak in the ISI density is caused by spikes occurring at the second maximum of the subthreshold oscillation. Therefore the distance between the first and the second peaks in the ISI density $\mathcal{F}(T)$ reflects the period of subthreshold oscillations. However, if no spike is generated within the relaxation time window $t_{\mathrm{rel}} = 2C/\gamma$ (i.e., not within the first few periods of subthreshold oscillations), their fixed phase relation to the last spike is lost due to the noise contribution. For this reason, the large intervals between spike clusters are random and lead to the observed exponential tail of $\mathcal{F}(T)$. The number of visible peaks in $\mathcal{F}(T)$ therefore depends on the ratio between the relaxation t_{rel} of subthreshold oscillations and their period. $\mathcal{F}(T)$ exhibits more peaks for large t_{rel} (i.e., when the effective damping coefficient γ/C is small).

With this insight, the different fit quality in Fig. 4*A* is easily explained: for small damping the first maximum of a subthreshold oscillation is rather large. To limit the number of action potentials triggered at this maximum, the spike threshold also needs to be high. This, however, implies that large intrinsic noise intensities are required to reach the experimentally measured firing-rate levels. Later on we show how this structural shortcoming of the current model can be overcome by a stochastic spike-generation mechanism.

For nonresonant cells, experimental ISI distributions were reproduced with ease (Fig. 4*B* and Table 1). The explanation is straightforward: for these neurons, the effective damping coefficient was relatively large, leading to a short relaxation time, which canceled the effects of any initial phase-locking; thus after each spike, the membrane voltage quickly returned to its steady-state value. Consequently, the position of the main peak in the ISI distribution was mainly determined by the noise intensity.

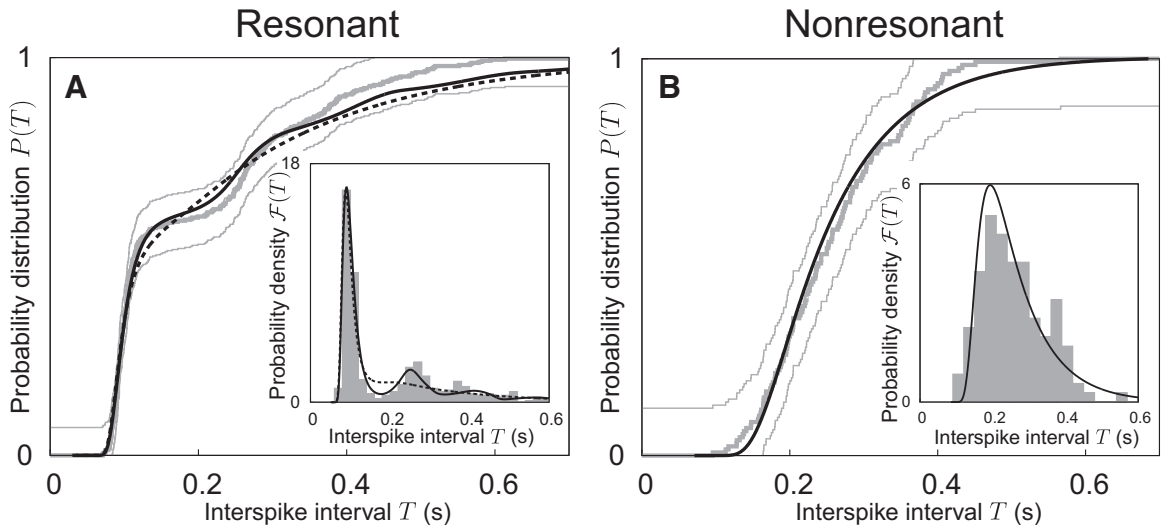


FIG. 4. ISI distributions. Predictions from the resonate-and-fire model are depicted as black lines, experimental results as thick gray lines, and thin gray lines mark the boundaries of a Kolmogorov–Smirnov test with significance level 5%. *Insets* depict the corresponding ISI densities. *A*: stellate cell, same as in Fig. 1. Two model variants with a different SD of the voltage fluctuations are shown: $\sigma_x = 5.6$ mV as solid line and $\sigma_x = 1.6$ mV as dashed line. The first variant provides the better fit to the measured data but its voltage fluctuations are much larger than experimental values ($\sigma_x \approx 2.5$ mV). The second variant shows the best fit obtained with an upper variance bound of 4.0 mV. *B*: pyramidal cell, same as in Fig. 1. The best approximation, shown in black, closely matches the experimental ISI distribution. The parameters of the underlying resonate-and-fire models are specified in Table 1.

Firing-rate–dependent changes in spike patterns and ISI distributions can also be reproduced within the resonate-and-fire framework. In stellate cells (Fig. 2*G*), the position of the first peak of $\mathcal{F}(T)$ did not vary with firing-rate increases and the tail of the density shortened. Because the subthreshold resonance frequency is fairly independent of the depolarization in stellate cells (Erchova et al. 2004; see also Fig. 3*C*), the model therefore suggests that the reset period τ_r also does not strongly change with depolarization. As the noise intensity grows with depolarization (Fig. 3*G*), the overall probability of spike generation increases. Consequently, the exponential tail of $\mathcal{F}(T)$ becomes shorter. In pyramidal cells, an increased firing rate simply shifts the position of the maximum in the ISI density toward shorter ISIs (Fig. 2*H*): the higher noise level increases the probability of spike generation and thus reduces the length of ISIs.

Higher-order ISI statistics

For resonant cells, a good agreement between model and experimental ISI distributions has, so far, been possible only at the cost of large voltage fluctuations. Because of the reset of all variables after each model spike, ISIs are independent of each other. The present model thus implements a renewal process. If, however, the assumption of independence of ISIs was incorrect, a modification of the model introducing ISI correlations through a nonrenewal mechanism would be required. This extension to the model may also further improve its overall performance. For this reason, we analyzed the experimental ISIs in more detail.

The probability density of the ISI return map $\mathcal{F}(T_1, T_2)$ quantifies the joint probability of obtaining two particular successive ISIs (Fig. 5, *A* and *B*). For stellate cells, the strong peak at about (100; 100) ms and the two orthogonal arms up to ISIs of about 500 ms reflect spike clusters. For pyramidal cells, the single broad maximum at about (250; 250) ms mirrors the unimodal ISI distribution (Fig. 5*B*). Both results agree with our previous findings. Next, we analyzed correlation maps that quantify the deviation of a measured ISI sequence from the ISI sequence of a renewal spike train with the same ISI distribution (Fig. 5, *C* and *D*). As revealed by this analysis, experimental ISIs are clearly not independent. To quantify the correlations between more distant spikes, serial correlation coefficients (SCCs) were used (Fig. 5, *E* and *F*). Here, the lag k denotes the distance between ISIs. Interestingly, negative SCC values for $k = 1$ (neighboring ISIs) were observed at higher firing rates (3–10 Hz) in both cell types. The correlation time τ_c ranged from 200 to 400 ms. For lower firing rates (<2 Hz), however, SCC values for all $k > 0$ did not significantly differ from zero. Finally, the number of correlated ISIs grew with increasing firing rate for both cell types (data not shown).

ISI correlations are captured by an extended resonate-and-fire model

The experimental correlations (Fig. 5, *C–F*) cannot be captured by a renewal model. We therefore included a spike-history–dependent stochastic firing threshold: after threshold crossing, spikes are triggered with only a certain probability; the occurrence of several spikes within a time window shorter than the correlation time τ_c lowers the probability to generate another spike and, consequently, not every threshold crossing elicits a spike (Fig. 6). The resulting nonrenewal model allows one to capture the observed ISI correlations as well as to match ISI distributions without excessive voltage fluctuations. As for the renewal model, parameters were optimized within the experimental bounds. Differences between experimental and model ISI distributions and absolute values of SCC (for $k = 1$) were now minimized simultaneously.

Although the experimental voltage fluctuations are no longer exceeded, the ISI distributions of the nonrenewal model agree better with the experimental data than the renewal-based fits (Fig. 7, *A* and *B*). This is because the nonrenewal threshold

mechanism limits the number of spikes generated in close succession. It is therefore possible to produce the required number of short intervals and simultaneously ensure a small damping coefficient γ/C and low threshold x_b. In addition, the serial correlation coefficients suggested by the model matched the experimentally obtained results for both cell classes (Fig. 7, *C* and *E*). We also established that the model is consistent with the observation that ISI correlations vanish at lower firing rates (Fig. 7, *D* and *F*). To that end, SCCs in the optimal nonrenewal model with decreased noise (both cell types) and increased damping coefficient (only stellate cells) were analyzed. Although the nonrenewal threshold mechanisms remained unchanged, no ISI correlations were observable at lower firing rates.

It is a property of both the renewal and nonrenewal models that the distance between the first and the second peaks of the ISI density corresponds to the period of subthreshold oscillations. As previously observed in the entorhinal cortex (Erchova et al. 2004) in accordance with the subthreshold part of the model, this oscillation period is also approximately equal to the inverse of the subthreshold resonance frequency. According to the model, the distance between the first and second peaks in the ISI density should therefore be approximately equal to the inverse of the subthreshold resonance frequency. Indeed, our

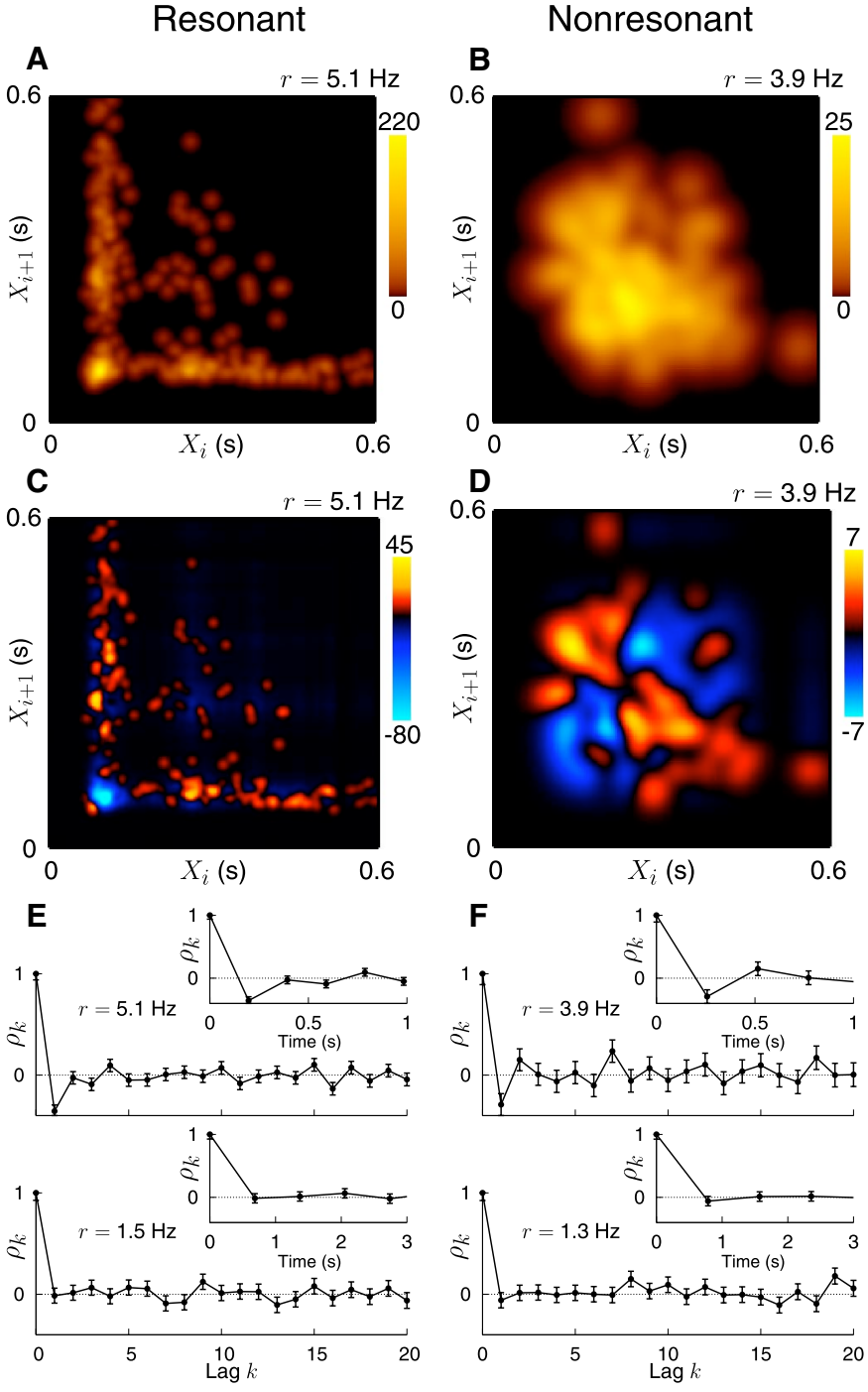


FIG. 5. Second-order ISI statistics for stellate (*left*) and pyramidal cells (*right*). Both neurons are the same neurons as in Fig. 1. *A* and *B*: probability density of the ISI return map. For stellate cells, spike clusters are reflected in the 2-arm structure with a pronounced peak. In contrast, the ISI return map of the pyramidal cell has a cloudlike structure, as expected for unclustered spike patterns. *C* and *D*: density of the ISI correlation maps. Both cell types exhibit negative correlations between adjacent ISIs. *E* and *F*: serial correlation coefficients (SCCs) ρ_k with lags $k = 1, \ldots, 20$ for 2 different average firing rates (stellate cell: 1.5 and 5.1 Hz; pyramidal cell: 1.3 and 3.9 Hz). The *insets* depict SCCs scaled by the respective mean ISI to represent the duration of correlations in seconds. SCCs of both stellate and pyramidal cells show similar negative correlations at short lags for sufficiently high firing rates. Independent of the cell type, the correlation time is about 300 ms.

experimental data confirm this prediction (Fig. 8). Across the electrophysiologically recorded set of resonant stellate cells (natural range of resonance frequency: 5–12 Hz; see Erchova et al. 2004), the inverse distance between the first two peaks in the ISI densities was found to be approximately equal to the measured subthreshold resonance frequency.

It is important to notice that the subthreshold resonance frequency in stellate cells is approximately constant with depolarization (Erchova et al. 2004). This observation implies that estimates of the resonance frequency at less depolarized levels of the membrane potential—and not only those measured close to threshold—correlate with the described firing statistics. In other cells or under different experimental conditions the subthreshold resonance frequency, however, can change with depolarization. In these cases, quantitative aspects of spike clustering may also depend on depolarization. For example, Nolan et al. (2007) reported that in stellate cells of HCN1 knockout mice, the probability of spikes occurring in clusters and the intracluster spike frequency were changed and resonance at the resting membrane potential was abolished. On the other hand, the observed intrinsic membrane-potential oscillations closely mimicked those in wild-type animals near the firing threshold. This indicates that an HCN1-independent slow potassium current is involved in spike clustering in these cells, in full agreement with our general model framework.

DISCUSSION

The parallel analysis of experimental and modeling data shows that ISI distributions of both stellate and pyramidal cells are strongly influenced by subthreshold processes. Membrane-potential resonance—an important signature of subthreshold dynamics—causes spike clustering. A low-dimensional renewal model constrained by measured experimental parameters reproduces the observed ISI distributions and allows us to quantitatively understand the mechanisms that lead to spike clustering. In particular, we observed that a spike-induced reset of subthreshold oscillations underlies the clustered spike patterns in resonant neurons. In addition, spike-induced processes cause ISI correlations *independently* of the existence of subthreshold resonance. A spike-history–dependent modification of the threshold mechanism in the model captures these higher-order ISI statistics.

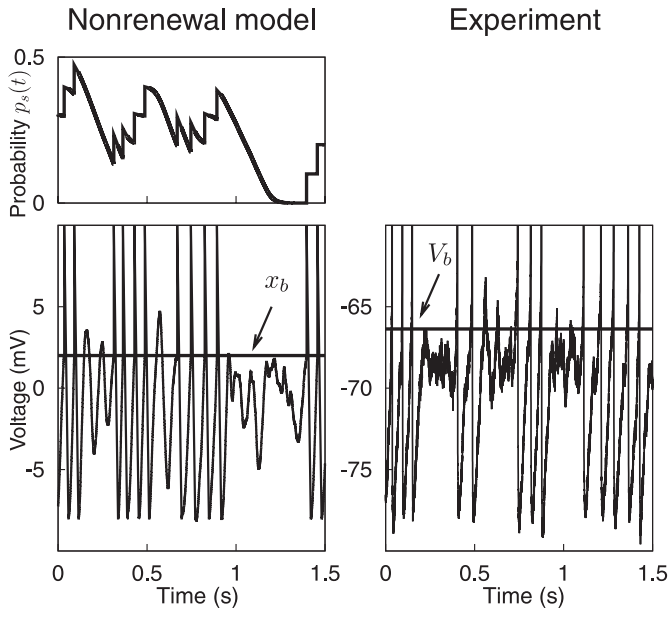


FIG. 6. Nonrenewal spike generation. *Left column*: example of a nonrenewal spike train generated by the extended resonate-and-fire model (*bottom*) and the corresponding time evolution of the probability to skip a spike $p_s(t)$ (*top*). In this model, the voltage variable may briefly exceed the threshold for spike generation without producing a spike, resulting in a "silent overshoot." *Right column*: experimental data from a stellate cell. Multiple silent overshoots are visible in this recording.

Subthreshold dynamics shape spike-train statistics

This study was performed at membrane depolarizations in the vicinity of the spiking threshold with average firing rates $\leq$10 Hz. With the majority of in vivo firing rates ranging from 0 to 10 Hz (Sargolini et al. 2006), the range of firing rates explored within our study thus covers the behaviorally relevant regime. Nevertheless, the cells are able to produce fast-spike sequences, in that the maximum instantaneous rates in both cell types were up to 20 Hz (ISIs of ~50 ms).

A relation between subthreshold resonance and clustered spike patterns has been conjectured previously for neurons in the olfactory bulb (Chen and Shepherd 1997; Desmaisons et al. 1999) and brain stem (Pedroarena et al. 1999; Wu et al. 2001). Our experimental and theoretical findings systematically show that in stellate cells subthreshold resonance is sufficient to explain spike clustering. No spike-induced activation of additional currents is necessary. In particular, the renewal model predicts that without subthreshold resonance no multimodal ISI distributions can be obtained.

Note that the proposed clustering mechanism in stellate cells is different from that of intrinsically bursting neurons. Although the latter also exhibit spike clusters at higher firing rates, the cells reported here fire regularly if sufficiently depolarized (Richardson et al. 2003; Schreiber et al. 2004b). Within the low-firing range, however, stellate cells in entorhinal cortex can potentially profit from the same computational advantages discussed for intrinsically bursting neurons. These include coding with intra- and intercluster time intervals (Kepecs and Lisman 2003), increase of reliability of synaptic transmission (Lisman 1997), and frequency-specific routing of information if ISIs within a cluster of presynaptic neurons match the membrane resonance of the postsynaptic cell (Izhikevich et al. 2003).

Implication of negative interspike-interval correlations

Negative ISI correlations have been reported for many different cell types, starting as early as Kuffler et al. (1957). Negative ISI correlations are observed even under constant stimulation (Ratnam and Nelson 2000) and it has recently been suggested that they enhance neuronal coding capabilities (Chacron et al. 2001). The correlations between ISIs observed here on a timescale of 200–400 ms are likely to be caused by slowly deactivating currents that are activated by the preceding spike. Such patterns, e.g., slow potassium currents, are known to mediate firing rate adaptation (Brown and Adams 1980; Madison and Nicoll 1984). Interestingly, these currents can also contribute to subthreshold resonance (Hutcheon and Yarom 2000). Our study, however, clearly shows that both

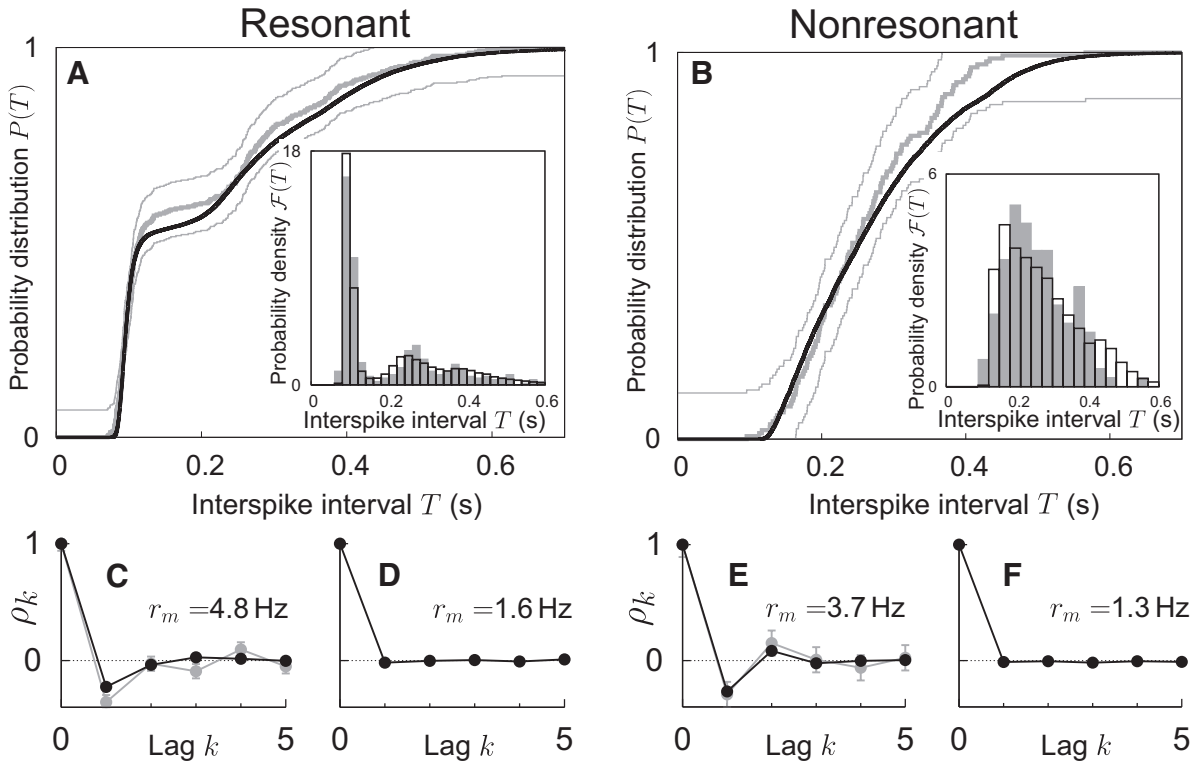


FIG. 7. Nonrenewal resonate-and-fire model explains ISI distributions and ISI correlations. *Left panels*: stellate cell; *right panels*: pyramidal cell (same cells as in Fig. 1). Experimental ISI statistics are shown in gray, modeling data in black. *A* and *B*: ISI distributions (main panels) and corresponding ISI densities (*insets*). *C–F*: serial correlation coefficients ρ_k. r_m specifies the mean firing rate in the model. *C* and *E* correspond to *A* and *B*, respectively. The values of the fit parameters are given in Table 1. *D* and *F* illustrate that the strong reduction of ISI correlations at lower firing rates (compare with *C* and *E*) can be reproduced in the model [$r_m = 1.6$ Hz for the stellate cell (*D*) with weaker noise $D = 0.003$ nA^2 Hz, larger friction $\gamma/C = 0.63 \times 10^2$ Hz and other parameters as in *A*; and $r_m = 1.3$ Hz for the pyramidal cell (*F*) with weaker noise $D = 0.003$ nA^2 Hz and other parameters as in *B*]. As shown by these results, even simple nonrenewal resonate-and-fire models can explain rather complex spike-train data.

phenomena can exist independently of each other: pyramidal cells exhibited negative correlations but did not show any subthreshold resonance.

The mechanistic insight gained from the nonrenewal model explains why the ISI correlations are visible only at higher firing rates in both cell types: spike-induced currents responsible for the temporal relations between spikes are likely to be active across a wide range of depolarizations. Nevertheless, these currents cannot influence spike patterns if the mean ISI is larger than the deactivation timescale. This is true at low firing rates. In contrast, for higher firing rates the mean ISI is sufficiently small for spike-induced currents to interact with spike generation. Along these lines, (*i*) mimicking a decreased firing rate in the optimal nonrenewal model made the negative correlations disappear (Fig. 7) and (*ii*) fits of experimental data with low firing rates were equally good with the renewal and nonrenewal models (data not shown).

Interestingly, for higher firing rates the nonrenewal spike mechanism also improved the representation of ISI distributions within the resonate-and-fire framework. This indicates that correlations between spikes are involved in shaping the fine structure of ISI distributions, whereas the fundamental features, such as spike clustering, are determined by the subthreshold frequency preference.

Model scope, limitations, and outlook

Why do we use a simple and transparent mathematical model for describing activity in cells of the entorhinal cortex? First, on a practical level, the model is analytically tractable and, at the same time, permits an accurate phenomenological description of both resonant and nonresonant neurons. Due to its generality it can be applied to different cell types, independently of a particular ionic realization of subthreshold resonance properties. Noticeably, the model parameters used to describe a given neuron are tightly constrained by experimental data and obtained from the very same cell. This is not possible for biophysically realistic models whose many parameters cannot be estimated with measurements from a single neuron.

Second, on a conceptual level, the current model provides a quantitative framework to investigate the relation between subthreshold and spiking activity. Here, neglecting microscopic details allows one to reveal general neural mechanisms that go far beyond any specific biophysical processes. This is important because the combinations of conductances may be vastly different between cells, but still lead to the same input–output transformation (Marder and Prinz 2002). In particular, the model clearly demonstrates that the well-known phenomenon of subthreshold resonance can be sufficient to fully explain clustered spike patterns and multimodal ISI distributions. Furthermore, it enables us to disentangle the contributions of subthreshold and spike-induced dynamics to the observed spike patterns. Although it is possible that specific ionic currents contribute to both spike-induced and resonance-related dynamics, our model shows that separating these dynamics on a phenomenological level still captures firing statistics.

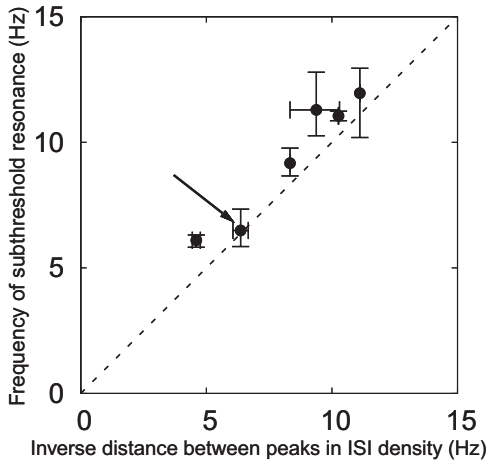


FIG. 8. Relation between the frequency of subthreshold resonance and the inverse distance between the first and second peaks in the ISI density. Each data point corresponds to a different stellate cell, which are known to have resonances in the frequency range of 5–15 Hz. The horizontal and vertical bars mark the minimal and maximal observed values, as described in detail in METHODS. The arrow points to the cell presented in Figs. 2–7. Our model framework predicts that for every cell the inverse differences between the two ISI peaks are very similar to the natural oscillation frequency near the firing threshold. This frequency, in turn, has been shown to be close to but smaller than the subthreshold resonance frequency (compare f_{osc} and f_{res}, *Eq. 4*), as discussed in METHODS and Erchova et al. (2004). It follows that the resonance frequency should slightly exceed the inverse ISI-peak difference so that the data points should be above but close to the dashed identity line, as confirmed by the population data.

In particular, describing spike-induced effects by a reset time and a spike-history–dependent spike probability results in a surprisingly good approximation of experimental data.

Interestingly, the mathematical model also suggests that the clustered spike patterns of EC stellate cells are obtained because each spike induces a phase reset of the intrinsic subthreshold oscillation so that the subsequent spike is likely to be generated after a fixed time lag, roughly equal to the reset period plus half an oscillation period. This picture differs from an alternative situation where spikes, such as those in the CA1 region, are preferentially generated without phase reset at the maxima of an ongoing subthreshold oscillation superimposed with a slowly varying excitatory input (see, e.g., Buzsáki 2002; Leibold et al. 2008). If that description also applied for spikes caused by the combination of intrinsic subthreshold oscillations and constant current injections, as investigated in the present study, the main peak of the intracluster ISI distribution should coincide with the subthreshold oscillation period. Our data show, however, that the dominant intracluster ISI is shorter than the oscillation period (Fig. 2). In contrast, the time between the first and second peaks is similar to the oscillation period, in agreement with the model.

We also point out that, according to the proposed model, the existence of persistent subthreshold oscillations relies on the presence of damped oscillations of the corresponding noiseless system ($D = 0$). For the observed oscillations to be visible, however, deviations from the rest state are required, which for synaptically blocked cells are triggered when (*i*) the cell is depolarized and (*ii*) intrinsic noise allows large deviations from the steady state that effectively translate into persistent oscillations. The difference to models like that of Fransén et al. (2004) and Rotstein et al. (2006) lies in the fact that persistent subthreshold oscillations in such a model can exist independently of noise. In agreement with our view that noise is needed to trigger the damped oscillations, Dorval and White (2005) have shown that channel noise from persistent sodium channels is essential for subthreshold membrane potential oscillations.

Within the current model framework, spike-induced changes in the voltage trajectory are described by a constant reset period and a fixed reset voltage. Spike-induced currents, such as those underlying afterhyperpolarization, may, however, be active longer than the model reset period and thus also contribute to the generation of later spikes. Other nonrenewal models either explicitly include slow spike-induced ionic currents (Benda and Herz 2003; Izhikevich 2004) or modify the threshold values dynamically (Chacron et al. 2004). These models, however, are not consistent with two observations in our experiments: First, the firing threshold was not higher for shorter than that for longer ISIs, as would be expected with a dynamic threshold. Second, the voltage sometimes crossed the threshold without producing spikes, resulting in "silent overshoots," as shown in Fig. 6. The stochastic threshold used in this study accounts for both observations and directly reflects channel noise at the spike-initiation zone. This noise is particularly high near the firing threshold and known to severely influence spike generation (Schneidman et al. 1998).

As demonstrated by our results, spike patterns of single EC stellate and pyramidal cells are shaped by a combination of subthreshold resonances and spike-induced ionic currents. The generality of the model suggests that similar mechanisms may also be relevant for other neural systems. It remains an open challenge—to be tackled in future studies—whether the signatures of resonance and nonresonance, respectively, are behaviorally relevant. Were this indeed the case, and recent results on grid cells point in that direction (Giocomo et al. 2007), information processing in the entorhinal cortex might be understandable on the basis of a simple mathematical framework. If, on the other hand, behaviorally relevant time-varying inputs generate responses that cannot be understood in terms of a resonate-and-fire mechanism, the interesting question would remain why stellate and pyramidal cells exhibit such strikingly different dynamics, both below and above firing threshold.

Findings from the present study support the first of these two alternatives: all our data suggest that both stellate and pyramidal cells respond only rather weakly to high-frequency inputs. In particular, we did not observe any resonance or strong intrinsic oscillation of the membrane potential above 20 Hz. Rapid synaptic input fluctuations, which might occur under in vivo conditions, would thus only marginally influence the cells's firing activity. On the other hand, slower changes of the mean membrane depolarization are likely to be encoded in the cell-type–specific output patterns of stellate and pyramidal neurons. In such a scenario, the current model framework offers a powerful tool for studying neural dynamics and signal processing under artificial as well as behaviorally relevant conditions. The simplicity of the model will also make it possible to readily extend those studies from the single neurons to the network level.

ACKNOWLEDGMENTS

We thank U. Heinemann and I. M. Sokolov for many stimulating discussions and M. Stemmler for comments on a previous version of the manuscript.

Present address of A. Herz: Department of Biology, Ludwig-Maximilians-Universität Munich and Bernstein Center for Computational Neuroscience, Munich, Germany.

GRANTS

This work was supported by Deutsche Forschungsgemeinschaft Grants SFB 555 and SFB 618, Bernstein Center for Computational Neuroscience Berlin, Biotechnology and Biological Sciences Research Council, Engineering and Physical Sciences Research Council, Medical Research Council/Doctoral Training Center at the University of Edinburgh, Lloyd Trustee Savings Bank Group plc, and a personal fellowship to I. Erchova from the Royal Society of Edinburgh.

REFERENCES

Alonso A, Klink R. Differential electroresponsiveness of stellate and pyramidal-like cells of medial entorhinal cortex layer II. *J Neurophysiol* 70: 128–143, 1993.

Benda J, Herz AVM. A universal model for spike-frequency adaptation. *Neural Comput* 15: 2523–2564, 2003.

Brown DA, Adams PR. Muscarinic suppression of a novel voltage-sensitive K^+ current in a vertebrate neuron. *Nature* 183: 673–676, 1980.

Buzsáki G. Theta oscillations in the hippocampus. *Neuron* 33: 325–340, 2002.

Buzsáki G. Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15: 827–840, 2005.

Buzsáki G, Draguhn A. Neuronal oscillations in cortical networks. *Science* 304: 1926–1929, 2004.

Chacron MJ, Lindner B, Longtin A. Noise shaping by interval correlations increases information transfer. *Phys Rev Lett* 21: 080601, 2004.

Chacron MJ, Longtin A, Maler L. Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. *J Neurosci* 21: 5328–5343, 2001.

Chen WR, Shepherd GM. Membrane and synaptic properties of mitral cells in slices of rat olfactory bulb. *Brain Res* 745: 189–196, 1997.

Desmaisons D, Vincent JD, Lledo PM. Control of action potential timing by intrinsic subthreshold oscillations in olfactory bulb output neurons. *J Neurosci* 19: 10727–10737, 1999.

Diba K, Lester HA, Koch C. Intrinsic noise in cultured hippocampal neurons: experiment and modeling. *J Neurosci* 24: 9723–9733, 2004.

Dorval ADJ, White JA. Channel noise is essential for perithreshold oscillations in entorhinal stellate neurons. *J Neurosci* 25: 10025–10028, 2005.

Erchova I, Kreck G, Heinemann U, Herz AVM. Dynamics of rat entorhinal cortex layer II and III cells: characteristics of membrane potential resonance at rest predict oscillation properties near threshold. *J Physiol* 560: 89–110, 2004.

Fransén E, Alonso AA, Dickson CT, Magistretti J, Hasselmo ME. Ionic mechanisms in the generation of subthreshold oscillations and action potential clustering in entorhinal layer II stellate neurons. *Hippocampus* 14: 368–384, 2004.

Geisler C, Brunel N, Wang XJ. Contributions of intrinsic membrane dynamics to fast network oscillations with irregular neuronal discharges. *J Neurophysiol* 94: 4344–4361, 2005.

Gimbarzevsky B, Miura RM, Puil E. Impedance profiles of perirhinal and central neurons. *Can J Physiol Pharmacol* 62: 460–462, 1984.

Giocomo LM, Zilli EA, Fransén E, Hasselmo ME. Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science* 315: 1719–1722, 2007.

Gutfreund Y, Yarom Y, Segev I. Subthreshold oscillations and resonant frequency in guinea-pig cortical neurons: physiology and modelling. *J Physiol* 483: 621–640, 1995.

Haas JS, White J. Frequency selectivity of layer II stellate cells in the medial entorhinal cortex. *J Neurophysiol* 88: 2422–2429, 2002.

Hutcheon B, Miura RM, Yarom Y, Puil E. Low-threshold calcium current and resonance in thalamic neurons: a model of frequency preference. *J Neurophysiol* 71: 583–594, 1994.

Hutcheon B, Yarom Y. Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neurosci* 23: 216–222, 2000.

Izhikevich EM. Resonate-and-fire neurons. *Neural Networks* 14: 883–894, 2001.

Izhikevich EM. Which model to use for cortical spiking neurons? *IEEE Trans Neural Networks* 15: 1063–1070, 2004.

Izhikevich EM, Desai NS, Walcott EC, Hoppensteadt FC. Bursts as a unit of neural information: selective communication via resonance. *Trends Neurosci* 26: 161–167, 2003.

Kepecs A, Lisman J. Information encoding and computation with spikes and bursts. *Network* 14: 103–118, 2003.

Klink R, Alonso A. Ionic mechanisms for the subthreshold oscillations and differential electroresponsiveness of medial entorhinal cortex layer II neurons. *J Neurophysiol* 70: 144–157, 1993.

Kuffler SW, Fitzhugh R, Barlow HB. Maintained activity in the cat's retina in light and darkness. *J Gen Physiol* 40: 683–702, 1957.

Ledermann W. *Handbook of Applicable Mathematics: Statistics*. Chichester, UK: Wiley, 1984, vol. 6.

Leibold C, Gundlfinger A, Schmidt R, Thurley K, Schmitz D, Kempter R. Temporal compression mediated by short-term synaptic plasticity. *Proc Natl Acad Sci USA* 105: 4417–4422, 2008.

Leung LS, Yu HW. Theta-frequency resonance in hippocampal CA1 neurons in vitro demonstrated by sinusoidal current injection. *J Neurophysiol* 79: 1592–1596, 1998.

Lisman J. The theta/gamma discrete phase code occurring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* 15: 913–922, 2005.

Lisman JE. Bursts as a unit of neural information: making unreliable synapses reliable. *Trends Neurosci* 20: 38–43, 1997.

Llinás R, Yarom Y. Oscillatory properties of guinea-pig inferior olivary neurones and their pharmacological modulation: an in vitro study. *J Physiol* 376: 163–182, 1986.

Madison DV, Nicoll RA. Control of the repetitive discharge of rat CA1 pyramidal neurones in vitro. *J Physiol* 354: 319–331, 1984.

Manwani A, Koch C. Detecting and estimating signals in noisy cable structures, I: neuronal noise sources. *Neural Comput* 11: 1797–1829, 1999.

Marder E, Prinz AA. Modeling stability in neuron and network function: the role of activity in homeostasis. *BioEssays* 24: 1145–1154, 2002.

Mauro A, Conti F, Dodge F, Schor R. Subthreshold behavior and phenomenological impedance of the squid giant axon. *J Gen Physiol* 55: 497–523, 1970.

Nolan MF, Dudman JT, Dodson PD, Santoro B. HCN1 channels control resting and active integrative properties of stellate cells from layer II of the entorhinal cortex. *J Neurosci* 27: 12440–12451, 2007.

Pedroarena CM, Pose IE, Yamuy J, Chase MH, Morales FR. Oscillatory membrane potential activity in the soma of a primary afferent neuron. *J Neurophysiol* 82: 1465–1476, 1999.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes in C*. Cambridge, UK: Cambridge Univ. Press, 1999.

Puil E, Gimbarzevsky B, Miura RM. Quantification of membrane properties of trigeminal root ganglion neurons in guinea pigs. *J Neurophysiol* 55: 995–1016, 1986.

Ratnam R, Nelson ME. Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals. *J Neurosci* 20: 6672–6683, 2000.

Richardson MJE, Brunel N, Hakim V. From subthreshold to firing-rate resonance. *J Neurophysiol* 89: 2538–2554, 2003.

Rotstein HG, Oppermann T, White JA, Kopell N. The dynamic structure underlying subthreshold oscillatory activity and the onset of spikes in a model of medial entorhinal cortex stellate cells. *J Comput Neurosci* 21: 271–292, 2006.

Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser MB, Moser EI. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* 312: 758–762, 2006.

Schaefer AT, Angelo K, Spors H, Margrie TW. Neuronal oscillations enhance stimulus discrimination by ensuring action potential precision. *PloS Biol* 4: 1010–1024, 2006.

Schneidman E, Freedman B, Segev I. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput* 10: 1679–1704, 1998.

Schreiber S, Erchova I, Heinemann U, Herz AVM. Subthreshold resonance explains the frequency-dependent integration of periodic as well as random stimuli in the entorhinal cortex. *J Neurophysiol* 92: 408–415, 2004a.

Schreiber S, Fellous JM, Tiesinga P, Sejnowski TJ. Influence of ionic conductances on spike timing reliability of cortical neurons for suprathreshold rhythmic inputs. *J Neurophysiol* 91: 194–205, 2004b.

Verechtchaguina T, Sokolov IM, Schimansky-Geier L. First passage time densities in non-Markovian models with subthreshold oscillations. *Europhys Lett* 73: 691–697, 2006a.

Verechtchaguina T, Sokolov IM, Schimansky-Geier L. First passage time densities in resonate-and-fire models. *Phys Rev E* 73: 031108, 2006b.

Verechtchaguina T, Sokolov IM, Schimansky-Geier L. Interspike interval densities of resonate and fire neurons. *Biosystems* 89: 63–68, 2007.

Villacorta JA, Panetsos F. Information coding by ensembles of resonant neurons. *Biol Cybern* 92: 339–347, 2005.

White J, Rubinstein J, Kay A. Channel noise in neurons. *Trends Neurosci* 23: 131–137, 2000.

White JA, Klink R, Alonso A, Kay AR. Noise from voltage-gated ion channels may influence neuronal dynamics in the entorhinal cortex. *J Neurophysiol* 80: 262–269, 1998.

Wu N, Hsiao CF, Chandler SH. Membrane resonance and subthreshold membrane oscillations in mesencephalic V neurons: participants in burst generation. *J Neurosci* 21: 3729–3739, 2001.

Behavioral/Systems/Cognitive

Same or Different? A Neural Circuit Mechanism of Similarity-Based Pattern Match Decision Making

Tatiana A. Engel and Xiao-Jing Wang
Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510

The ability to judge whether sensory stimuli match an internally represented pattern is central to many brain functions. To elucidate the underlying mechanism, we developed a neural circuit model for match/nonmatch decision making. At the core of this model is a "comparison circuit" consisting of two distinct neural populations: match enhancement cells show higher firing response for a match than a nonmatch to the target pattern, and match suppression cells exhibit the opposite trend. We propose that these two neural pools emerge from inhibition-dominated recurrent dynamics and heterogeneous top–down excitation from a working memory circuit. A downstream system learns, through plastic synapses, to extract the necessary information to make match/nonmatch decisions. The model accounts for key physiological observations from behaving monkeys in delayed match-to-sample experiments, including tasks that require more than simple feature match (e.g., when BB in ABBA sequence must be ignored). A testable prediction is that magnitudes of match enhancement and suppression neural signals are parametrically tuned to the similarity between compared patterns. Furthermore, the same neural signals from the comparison circuit can be used differently in the decision process for different stimulus statistics or tasks; reward-dependent synaptic plasticity enables decision neurons to flexibly adjust the readout scheme to task demands, whereby the most informative neural signals have the highest impact on the decision.

Introduction

Perception and cognition often require us to evaluate similarity of two sensory events and to judge whether they are the same or different. "Same versus different" comparison is a generic neural computation involved in a wide range of brain functions. For example, searching for an object in a crowded scene requires us to judge whether a currently viewed object matches an internal representation of the target object. Furthermore, mismatch between expected and experienced stimuli is believed to give rise to "prediction error" signals [e.g., in the forward model for motor learning (Wolpert and Miall, 1996)]. Match/nonmatch comparison between the environment and expectation has also been proposed to gate the entry of information into the long-term memory (Lisman and Grace, 2005; Kumaran and Maguire, 2007).

Match/nonmatch computation is often thought of as a decision on whether the difference between two signals is zero (match). However, recent experimental findings in delayed match-to-sample (DMS) tasks suggest a different view. In a DMS task, subjects are presented with a sequence of stimuli separated by delays, and a behavioral response is required if the current test stimulus is the same (match) as a previously shown sample stimulus. Intriguingly, converging evidence from physiological studies with behaving monkeys and human brain imaging (Turk-Browne et al., 2007; Duncan et al., 2009) pointed to two candidate neural mechanisms involved in match versus nonmatch computation. One is referred to as repetition suppression, a passive reduction of neural response to any stimulus repetition regardless of behavioral relevance (see Fig. 1*C*). Repetition suppression is the predominant neural signal observed in standard DMS tasks (see Fig. 1*B*) when the matching test is the only stimulus repetition within a trial (Miller et al., 1991, 1993; Miller and Desimone, 1994; Steinmetz and Constantinidis, 1995; Constantinidis and Steinmetz, 2001; Zaksas and Pasternak, 2006). The other is referred to as match enhancement, an active mechanism that is engaged whenever feature matching is not sufficient to perform a task, as for example when nonmatch can also be repetitive (e.g., ABBA) (see Fig. 1*B*), and irrelevant repetitions of nonmatch stimuli (BB) should be ignored. Neurophysiological recordings in the prefrontal (Miller et al., 1996; Freedman et al., 2003), temporal (Miller and Desimone, 1994), and parietal (Rawley and Constantinidis, 2010) cortices revealed two populations of neurons whose selectivity for visual stimuli is modulated by match/nonmatch context in complementary ways: match enhancement (ME) cells show higher firing response for a match than nonmatch to the sample, whereas match suppression (MS) cells exhibit the opposite trend (see Fig. 1*D*).

These observations raised a number of questions: (1) what are the network mechanisms for generating match enhancement and suppression neural signals, (2) how does the brain switch between the active and passive modes of computation, and (3) are enhancement and suppression neural signals sufficient to make same versus different decisions, and if so, how? Here, we examine

Received Nov. 24, 2010; revised March 5, 2011; accepted March 13, 2011.

Author contributions: T.A.E. and X.-J.W. designed research; T.A.E. performed research; T.A.E. analyzed data; T.A.E. and X.-J.W. wrote the paper.

This work was supported by National Institutes of Health Grant R01 MH062349, The Swartz Foundation, and The Kavli Foundation. We thank S. Ardid and J. D. Murray for useful discussions.

Correspondence should be addressed to Xiao-Jing Wang, Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510. E-mail: xjwang@yale.edu.

DOI:10.1523/JNEUROSCI.6150-10.2011

 0270-6474/11/316982-15$15.00/0

possible answers to these questions by proposing a biophysically based circuit model that can learn and perform a DMS task in its entirety.

Materials and Methods

For the sake of concreteness, model simulations were performed with a DMS task in which the stimulus feature is the direction of motion in a field of moving dots (see Fig. 1*A*). Using the motion direction stimuli has three main advantages. First, the angle separation between any two motion directions is an analog quantity that objectively measures their similarity. Parametrical variation of the angle between the sample and test directions allows us to make quantitative predictions about neural encoding of similarity. Second, in the primates, processing of motion directions depends on neural activity in the cortical area MT, where most cells have bell-shaped tuning functions for the direction of motion (Dubner and Zeki, 1971; Britten et al., 1992; Born and Bradley, 2005). The encoding of motion directions by MT neurons is understood fairly well and can be captured with recurrent neural network models. Finally, the behavioral consequences of task difficulty for learning and performance can be studied by varying the fraction of test stimuli that are similar (less discriminable) to the sample. Although in this paper we focus on motion directions, all presented computational principles are generic and can be applied to other types of stimulus patterns.

Description of the model. The model consists of three interconnected subsystems: the working memory (WM), comparison, and decision networks. All three are strongly recurrent networks with dynamics governed by local excitation and feedback inhibition (Compte et al., 2000; Wang, 2002; Wong and Wang, 2006). In simulations, we used a reduced firing-rate model that has been shown to reproduce neural activity of a full spiking neuron network (Wong and Wang, 2006). In this framework, the dynamics of each excitatory neural pool is described by a single variable s representing the fraction of activated NMDA conductance, and the neural firing rate is described as a function of the total synaptic current. The variable s is described by the following:

$$\frac{ds}{dt} = -s/\tau_s + (1 - s)\gamma f(I), \tag{1}$$

with $\gamma = 0.641$ and $\tau_s = 60$ ms. The firing rate r is a function of the total synaptic current I (Abbott and Chance, 2005; Wong and Wang, 2006) as follows:

$$r = f(I) = \frac{aI - b}{1 - \exp[-d(aI - b)]}, \tag{2}$$

with $a = 270$ Hz/nA, $b = 108$ Hz, $d = 0.154$ s. The total synaptic current I consists of three main contributions: recurrent, sensory, and noisy, $I = I_r + I_s + I_n$. Recurrent input to a neuron i in the population A originating from the population B reads as follows:

$$I_{r,i}^{B\to A} = \sum_{j\in B} g_{ij}^{B\to A} S_j^B, \tag{3}$$

where $g_{ij}^{B\to A}$ is a synaptic coupling between the neuron j in the population B and the neuron i in the population A.

Neurons in the WM and comparison networks are spatially organized and labeled by their preferred direction of motion θ_i (from 0° to 360°). Each population (WM, ME, and MS) was simulated by 256 discrete units s_i ($i = 1 \ldots 256$) with equally spaced preferred directions ($\theta_{i+1} - \theta_i = 360°/256$). Within each network, the synaptic couplings g_{ij} between neurons with preferred directions θ_i and θ_j have a Gaussian profile as follows:

$$g_{ij}(\theta_i - \theta_j) = J_- + J_+ \exp(-(\theta_i - \theta_j)^2/2\sigma^2), \tag{4}$$

with $\sigma = 43.2°$. Parameters J_- and J_+ determine the amount of the recurrent inhibition and excitation in the circuit. The WM network can sustain persistent firing by reverberating activity because of strong recurrent excitation ($J_+^{WM} = 2.2$ nA, $J_-^{WM} = -0.5$ nA). In Figure 6, the peak location of persistent activity pattern was characterized by a population vector (Compte et al., 2000).

The comparison network has match enhancement and suppression (ME and MS) neurons defined by heterogeneous top–down inputs. One subpopulation (ME neurons) receives excitation from the WM circuit with the Gaussian profile as in Equation 4 and $\sigma = 43.2°$, $J_-^{WM\to ME} = 0$ nA, $J_+^{WM\to ME} = 1.15$ nA. The rest of the comparison network are MS neurons that do not receive any top–down input $J_-^{WM\to MS} = J_+^{WM\to MS} = 0$ nA. We assume that excitatory conductances of the ME cells are weakened by a factor $\alpha = 0.975$ because of a homeostatic mechanism acting to compensate for the top–down excitation in these cells. This homeostatic mechanism is operating on a very slow timescale, so that the value of α is held constant in all simulations. The comparison network is strongly dominated by inhibition with $J_- = -8.5$ nA, $J_+^{MS\to MS} = J_+^{ME\to MS} = 0.4$ nA and $J_+^{ME\to ME} = J_+^{MS\to ME} = \alpha J_+^{MS\to MS}$.

When a motion direction stimulus θ_s is presented, neurons in the WM and comparison networks receive sensory currents that depend on the preferred direction θ of the neuron as follows:

$$I_s = g_s \exp(-(\theta_s - \theta)^2/2\sigma_s^2), \tag{5}$$

where $\sigma_s = 43.2°$, $g_s^{WM} = 0.02$ nA, $g_s^{MS} = 0.13$ nA, and $g_s^{ME} = \alpha g_s^{MS}$. We assume that sensory signals reach the WM circuit only when attention is directed to store the sample in the WM. Signals form the test stimuli, as well as from the sample in the passive condition (simulating the repetition suppression) do not reach the WM circuit. In all simulations, sensory stimuli were presented for 0.6 s and separated by 1 s delay (except for the results in Fig. 6).

Noisy current replicates background synaptic inputs and obeys: $\tau_n dI_n/dt = -(I_n - I_0) + \sqrt{\tau_n}\sigma_n \eta(t)$, where $\eta(t)$ is a white Gaussian noise, $I_0^{MS} = 3.1$ *nA*, $I_0^{ME} = \alpha I_0^{MS}$, $I_0^{WM} = 0.3297$, $\tau_n = 2$ ms, and $\sigma_n = 0.009$ nA. For the results in Figure 6, the noise variance in the WM circuit was increased to $\sigma_n^{WM} = 0.16$ nA.

The ME and MS neurons have an additional current I_a mimicking the spike rate adaptation as follows: $I^{\{ME,MS\}} = I_r + I_s + I_n + I_a$, whereby $I_a = g_a s_a$ and $g_a = 0.003$ nA. The dynamics of s_a follows $ds_a/dt = -s_a/\tau_a + r$, with $\tau_a = 10$ s. We used a phenomenological model for the adaptation current, since our aim was to explore interactions between the passive and active memory mechanisms rather than to capture the precise biophysical mechanism of adaptation.

The strength of the top–down connections $J_+^{WM\to ME}$ and the homeostatic scaling parameter α were chosen so as to (1) achieve approximately equal responses in the ME cells to the preferred match and in the MS cells to the preferred nonmatch stimulus, and (2) replicate the experimentally observed difference in response to the match and nonmatch stimuli in the MS cells (see Fig. 1*D*). The magnitude of the adaptation current g_a was adjusted to mimic the experimental pattern of the passive repetition suppression in the MS cells (see Fig. 1*C*). Other observed firing rate patterns in the comparison network (as discussed in Results) were not purposely tuned.

The activities of the ME and MS neurons are pooled by the decision circuit with two competing neural populations selective for choice "match" and "nonmatch" (see Fig. 1*E*). When stimulated, activities of the two populations diverge according to winner-take-all dynamics, and the decision of the model is determined by the population with a higher activity. Across trials, the stochastic choice behavior of the decision circuit is characterized by a sigmoidal dependence of the probability P^M to choose match on the difference ΔI in synaptic input currents to the match and nonmatch pools (Soltani and Wang, 2006):

$$P^M = (1 + \exp(-\beta \Delta I))^{-1}. \tag{6}$$

We used $\beta = 200$ nA^{-1}.

Plasticity rule. The synapses connecting comparison neurons with the decision neurons are plastic. Each pair of presynaptic and postsynaptic cells is connected by a set of binary synapses that are in either a potentiated or a depressed state. The fraction c_{pre}^{post} of synapses in the potentiated state quantifies the strength of synaptic connection. Input currents to the match and nonmatch populations are expressed through the synaptic strengths as $I^{\{M,NM\}} = g\Sigma_i c_i^{\{M,NM\}} r_i$, where the

sum goes through all neurons in the comparison network, r_i are their firing rates, and $g = 1$ nA/Hz.

At the end of each trial, all synapses onto the chosen population (match or nonmatch) are updated according to a reward-dependent Hebbian plasticity rule. If the choice of the model is rewarded, the synapses are potentiated [i.e., the synapses in the depressed state make a transition to the potentiated state with the rate $q_0 \cdot q(r)$ referred to as the learning rate (Fusi, 2002)] as follows:

$$c \rightarrow c + q_0 \cdot q(r)(1 - c). \quad (7)$$

If the choice of the model is not rewarded, the synapses are depressed as follows:

$$c \rightarrow c - q_0 \cdot q(r)c. \quad (8)$$

The maximal learning rate q_0 determines the speed of learning. The learning rate gradually depends on the presynaptic firing rate: $q(r) = (1 + \exp(-(r - r_0)/\sigma_q))^{-1}$. We used $r_0 = 15$ Hz and $\sigma_q = 4$ Hz. For the results in Figure 9, we used $q_0 = 10^{-3}$.

Simulations of the learning dynamics. For modeling the learning process, it is computationally impractical to simulate the actual neural circuit (operating on the timescale of milliseconds) over thousands of trials. We devised the following approach to bypass this difficulty while faithfully capturing the behavior of the system. First, for the decision network, only the choice behavior but not the detailed temporal dynamics is important for learning. Therefore, on each trial, we evaluated the difference in the input currents ΔI, computed P^M using Equation 6, and then flipped a biased coin to determine the network choice on a single trial.

Second, we note that responses of the comparison neurons are not affected by learning, which only adjusts the readout scheme from these neurons. Therefore, to efficiently simulate the learning dynamics, we created a database of neural responses to different combinations of sample and test stimuli and used the database to investigate the learning process. Specifically, for each stimulus configuration, 100 trials of the model dynamics were simulated and stored in the database (except for 500 trials were simulated for the results in Fig. 6). Each trial in the simulations of the learning dynamics consisted of four sequential steps: (1) generate the sample and test motion directions according to the stimulus statistics; (2) choose one trial from the database that corresponds to the current sample and test directions; (3) evaluate the input currents to the decision circuit $\Delta I = g\Sigma_i(c_i^M - c_i^{NM})r_i$, and determine the network choice; (4) update synapses according to the learning rule. This approach is very efficient, since the database needs to be created only once, and then learning dynamics can be simulated for different stimulus statistics and different parameters of the plasticity rule using the same database.

The synaptic strengths were initialized with random values drawn from the uniform distribution on [0, 1]. After the learning dynamics reached a steady performance level, the psychometric function was obtained by averaging the model performance over 10^6 trials (with ongoing learning).

Steady-state calculations of the model performance. The synaptic strengths and model performance in the steady state can be calculated analytically. Let θ_s be the sample direction, which is uniformly distributed on [0°, 360°]. Possible directional differences, match $\theta_0 = 0°$ and nonmatch $\{\theta_i \neq 0°\}$ $(i = 1 \ldots N)$, have the priors p_0 and $\{p_i\}$, respectively. The firing rates of neurons in the comparison network $r_i(\theta,\theta_s)$ depend on the preferred direction θ, sample direction θ_s, and the directional difference θ_i of the neuron. Hence the learning rate of each neuron on every trial $q[r_i(\theta,\theta_s)]$ also depends on θ, θ_s, and θ_i. Averaging the update rule Equations 7 and 8 over the sample direction results in the effective learning rate as follows:

$$q_i(\theta) = (360°)^{-1}\int_{0°}^{360°} q[r_i(\theta,\theta_s)]d\theta_s.$$

The effective learning rate is different for ME and MS neurons because of difference in their firing rates, but it is the same for neurons with all preferred directions θ because of rotational symmetry of the ring architecture. Consequently, two sets of effective learning rates $\{q_i^{ME}\}$ and $\{q_i^{MS}\}$ determine the steady state of learning (index i refers to the directional difference θ_i).

Since the effective learning rate does not depend on θ, the steady-state values of synaptic strengths are also the same for neurons with all preferred directions. Hence, four synaptic strengths fully characterize the steady state as follows: c_{ME}^M, c_{ME}^{NM}, c_{MS}^M, c_{MS}^{NM}. The synaptic strengths of ME and MS neurons obey the same equations, but they differ because of different effective learning rates. The analytical expressions for the synaptic strengths are readily obtained as follows:

$$c^M = \frac{p_0 P_0^M q_0}{p_0 P_0^M q_0 + \sum_{i=1}^{N} p_i P_i^M q_i}, \quad (9)$$

$$c^{NM} = \frac{\sum_{i=1}^{N} p_i(1 - P_i^M)q_i}{p_0(1 - P_0^M)q_0 + \sum_{i=1}^{N} p_i(1 - P_i^M)q_i}. \quad (10)$$

Here P_i^M denotes the probability to choose match when the ith directional difference is presented. The difference in synaptic strengths to the match and nonmatch populations $\Delta c = c^M - c^{NM}$ determines the difference in the synaptic input currents $\Delta I_i = g[\Delta c_{ME}\Sigma_\theta r_i^{ME}(\theta) + \Delta c_{MS}\Sigma_\theta r_i^{MS}(\theta)]$, which in turn determines P_i^M. Hence Equations 9 and 10 have to be solved self-consistently, and we solve them numerically using the Levenberg–Marquardt algorithm.

Once the steady-state solution is obtained, P_i^M provides us the psychometric function. The overall performance (i.e., the overall fraction of correct responses) is then computed as follows: $p_0 P_0^M + \Sigma_{i=1}^N p_i(1 - P_i^M)$. The psychometric functions have sigmoidal shape and can be fitted with the function $f(\theta) = c/(1 + \exp(b(\theta - a)))$ of three parameters a, b, and c. The fitted value of $cb/4$ (measured in degrees^{-1}) is called the slope of the psychometric function and characterizes its steepness. The parameter c is the value of the psychometric function at 0° directional difference (i.e., represents the probability to correctly identify match). The psychometric threshold is defined as the sample-test directional difference at which the performance is 75% correct responses and is expressed through the fit parameters as $a - b\log(4c - 1)$.

In our model, the steady-state values of the synaptic strengths (Eqs. 9, 10) depend on the prior probabilities for match p_0 and nonmatch stimuli $\{p_i\}$. In this way, the model adjusts the behavioral output to various stimulus statistics, for example, when the match prior p_0 changes. Notably, the network model is not explicitly provided with the priors but learns them through experience.

Ideal Bayesian observer. As a benchmark against which to evaluate the network performance, we consider an ideal observer that performs the task optimally using Bayesian inference. On each trial, the ideal observer makes a match versus nonmatch decision based on observed data x (e.g., the firing rate) and the knowledge of priors p_0, $\{p_i\}$. Let $p(x|\theta_i)$ denote the likelihood function of x when the directional difference θ_i is presented. The posterior distributions for match and nonmatch are computed using Bayes' rule:

$$p(\text{match}|x) = \frac{p(x|\theta_0)p_0}{p(x|\theta_0)p_0 + \sum_{i=1}^{N} p(x|\theta_i)p_i}. \quad (11)$$

where the denominator is $p(x)$, and $p(\text{nonmatch}|x) = 1 - p(\text{match}|x)$. These posterior distributions can be used to make a decision using one of several possible decision rules. For the strict Bayesian strategy, the alternative with the larger posterior is always selected; hence the probability to choose match equals the following:

$$P(\text{match choice}|x) = \begin{cases} 1, & \text{if } p(\text{match}|x) > 0.5, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

For a probabilistic Bayesian strategy, the alternatives are chosen with probabilities equal to their posteriors; hence $P(\text{match choice}|x) = p(\text{match}|x)$ in this case. The psychometric function P_i^M for the ideal ob-

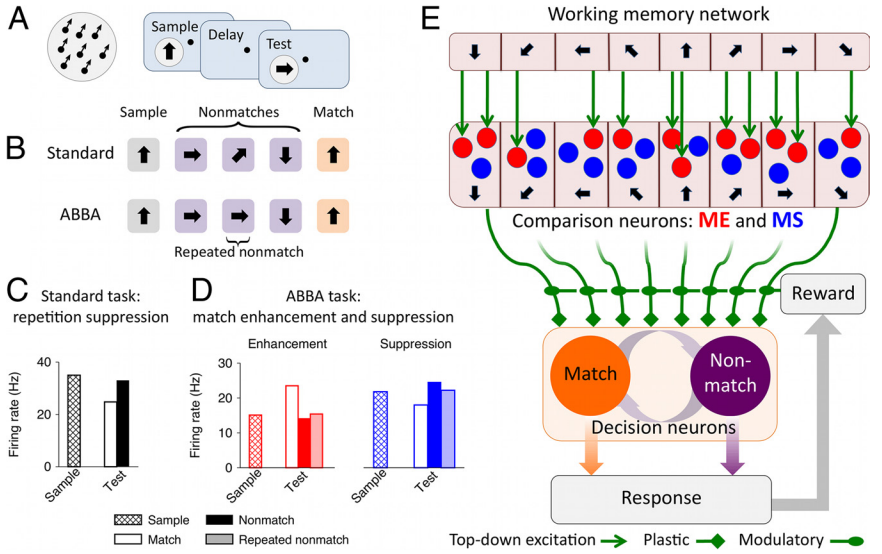


Figure 1. Delayed match-to-sample task, neural encoding of match/nonmatch, and schematic of the circuit model. ***A***, Left, Random dot stimulus. Right, The DMS task. The sample stimulus is followed by a sequence of test stimuli separated by delays. A behavioral response is required if the test matches the sample. ***B***, Example trials in two versions of the DMS task. In the standard task, all intervening nonmatches are different, and the match is the only perceptual stimulus repetition within a trial. In the ABBA task, irrelevant repetitions of nonmatches should be ignored. ***C***, Repetition suppression in inferior temporal cortex neurons in the standard DMS task [data from Miller et al. (1993)]. Average responses across cells to the same set of stimuli appearing as a sample, match, and nonmatch. ***D***, Match enhancement and match suppression in two complementary populations of prefrontal cortex neurons in the ABBA task [data from Miller et al. (1996)]. Average responses across cells to the same set of stimuli appearing as a sample, match, nonmatch, and repeated nonmatch. ***E***, Schematic of the circuit model. Neurons in the WM and comparison networks (ME and MS subpopulations) are tuned to directions of motion (indicated by arrows) and receive directional bottom–up input. Top–down projections from the WM to the comparison network are heterogeneous. ME neurons (red circles) receive stronger top–down excitation than MS neurons (blue circles). The decision network (match and nonmatch subpopulations) generates categorical match versus nonmatch choices by pooling activities of the ME and MS neurons through synapses that undergo reward-dependent Hebbian plasticity.

server is then computed for each directional difference θ_i by averaging $P(\text{match choice}|x)$ over the probability to observe the data x:

$$P_i^{\text{M}} = \int P(\text{match choice}|x)\, p(x|\theta_i)\, dx. \quad (13)$$

We assumed that, on each trial, the observed data value is $x = r(\theta_i) + \eta$, where $r(\theta_i)$ is the mean response when the directional difference θ_i is presented, and η is a Gaussian noise with zero mean and SD σ. Hence the likelihood $p(x|\theta_i) = (\sqrt{2\pi\sigma})^{-1}\exp(-(x - r(\theta_i))^2/(2\sigma^2))$. We considered two different choices for $r(\theta_i)$: (1) average firing rate of the ME population $[r_{\text{ME}}(\theta_i)]$ (see Fig. 3*C*, red line); (2) difference in the average firing rates of the ME and MS populations $[r_{\text{ME}}(\theta_i) - r_{\text{MS}}(\theta_i)]$ (see Fig. 3*C*, red and blue lines). We also considered the case when x is a two-component vector with the mean $\{r_{\text{ME}}(\theta_i), r_{\text{MS}}(\theta_i)\}$ and with two independent Gaussian noises. Performance of the ideal observer was very similar in all these cases and for both, strict Bayesian and probabilistic Bayesian, decision strategies.

For the comparison with the network model in Figure 7, we computed performance of the ideal observer using x as the difference in the average firing rates of the ME and MS populations and the strict Bayesian strategy. The noise SD σ was adjusted to approximately match the psychometric threshold and the overall performance for the network model and the ideal observer for $p_0 = 0.5$.

Alternative model. In the core of our model (see Fig. 10*A*, two-pool comparison model) are two neural populations, ME and MS neurons, that perform the comparison computation and exhibit complementary tuning to the sample test similarity (see Fig. 3*C*). We have also considered an alternative model based on simple addition of two signals: sensory input from the test stimulus and WM input representing a stored sample. The addition computation can be performed by a single neural population with converging sensory and WM inputs (see Fig. 10*A*, one-pool addition model). We implemented the one-pool addition model similarly to our two-pool model; however, instead of heterogeneous (ME and MS) comparison population, the one-pool model has a single "addition population" that receives sensory and WM inputs. All neurons in the addition population receive excitation from the WM circuit with the Gaussian profile as in Equation 4 and $J_+ = 1.15$ nA and $J_- = 0$ nA. Since the top–down excitation is homogeneous, there is no heterogeneity in the strengths of recurrent connections, bottom–up inputs and background noisy currents within the addition population. For all cells, we set $g_s = 0.13$ nA, $I_0 = 3.1$ nA, and the recurrent connections follow the Gaussian profile (Eq. 4) with $J_+ = 0.4$ nA and $J_- = -8.5$ nA. Other parameters are the same as in the two-pool model.

In the one-pool model, larger overlap between the top–down and bottom–up inputs leads to higher overall activity in the addition population. As a result, the average firing rate in the addition population gradually decreases with directional difference between the sample and test, resembling similarity tuning of the ME neurons in the two-pool model (see Fig. 10*B*,*C*, solid black lines). Match/nonmatch decisions can be read out from the single addition population by a simple threshold mechanism. Specifically, we assumed that the probability of the match decision is given by a sigmoidal function as follows:

$$P^{\text{M}} = \left(1 + \exp\left(-\frac{r - r_{\text{th}}}{\sigma_r}\right)\right)^{-1}, \quad (14)$$

where r is the averaged firing rate in the addition population, r_{th} is the firing-rate threshold, and parameter σ_r determines precision of the readout system.

To illustrate differences in behavioral performance of the one- and two-pool models, we asked how robust is the performance of each model to changes in the input strength (e.g., because of change in the contrast of visual stimuli) (see Fig. 10*A*,*B*). To this end, we simulated neural activity in both models in response to test stimuli with control ($g_s = 0.13$ nA) and doubled ($g_s = 0.26$ nA) strength. For fair comparison, with the control stimulus strength, the parameters r_{th} and σ_r in Equation 14 for the one-pool model were adjusted such that the psychometric function matches for the two models. With the doubled stimulus strength, the performance of both models was tested with the parameters of the readout systems fixed at the values obtained for the control stimulus strength.

Results

Computational hypotheses: building blocks of the circuit model

The model comprises three interconnected local circuits that correspond to three basic operations involved in the DMS task: the WM, comparison, and decision neural networks (Fig. 1*E*) (for details, see Materials and Methods). Neurons in the WM and comparison networks are tuned to motion directions and receive directional bottom–up inputs. The top–down projections from the WM circuit to the comparison network are excitatory and topographically organized: neurons with similar preferred directions are more strongly connected. Sample stimulus triggers persistent firing in the WM circuit, which represents a memory of the sample. This internal representation of the sample is maintained during the delay through reverberating neural activity (Camperi and Wang, 1998; Compte et al., 2000; Gutkin et al., 2001; Wang,

2001) and provides a top–down signal to modulate neural responses to test stimuli in the comparison network.

The core component of the model is the comparison network. Neurons in the comparison network respond differently to the test stimuli depending on whether they match the sample and in this way implement the comparison operation. The match/nonmatch sensitive modulations of responses arise from three simple biophysical ingredients. First, all cells in the comparison network are endowed with an adaptation current with a long time constant (~10 s) (Sanchez-Vives et al., 2000; Wang et al., 2003; Pulver and Griffith, 2010). The spike rate adaptation leads to a diminished response to any repeated stimulus and thus captures passive repetition suppression. Second, the top–down projections from the WM circuit are topographically organized but naturally heterogeneous: just by chance different cells within each column receive different amount of top–down excitation. The cells that receive stronger top–down excitation (Fig. 1*E*, red) show active ME, and the cells that receive weaker top–down excitation (Fig. 1*E*, blue) show MS, as explained in the following section. Finally, homeostatic regulation of excitatory synapses (Turrigiano et al., 1998; Renart et al., 2003) acts to maintain the average firing rate in the network and to keep the overall amount of excitation approximately equal for all cells. As a result, the recurrent and bottom–up synapses on the ME cells are slightly weakened to compensate for the top–down excitation, compared with the MS cells. As we shall see, the difference in strength of recurrent connections in the ME and MS cells is crucial to generate enhanced responses to nonmatches in the MS cells. Note that the homeostatic mechanism is operating on a very slow timescale; hence in all simulations the difference in strength of recurrent connections in the ME and MS cells is held constant.

It is noteworthy that the model assumes that the ME and MS effects arise naturally from heterogeneous top–down excitation and inhibition-dominated recurrent dynamics in the comparison network, and no learning is involved in shaping responses of the ME and MS neurons. It is possible that different tasks may engage ME and MS cells differently. For instance, in a task in which working memory might not be necessary, the ME cells might not receive top–down inputs and therefore would show passive repetition suppression.

The activity of the ME and MS neurons is readout by a downstream decision network, modeled similarly as in the previous work (Wang, 2002; Wong and Wang, 2006), that generates categorical match versus nonmatch decisions. The decision network comprises two neural populations: match neurons (Fig. 1*E*, orange) and nonmatch neurons (Fig. 1*E*, purple) fire at higher rate for match and nonmatch decisions, respectively. Unlike the comparison neurons, which exhibit ME and MS as a modulation of their selectivity for motion direction, the decision neurons carry a pure decision (response) signal and are not selective to any stimulus feature. In addition, the decision neurons acquire their decision (response) preferences through learning. The synapses connecting the comparison and decision networks undergo reward-dependent Hebbian plasticity (Soltani and Wang, 2006, 2010; Fusi et al., 2007). We will show that learning ultimately generates connectivity profiles such that the activity of the ME and MS neurons can be read out differently by the decision network in a way that allows flexible mapping of comparison signals onto arbitrary motor response. The model is able to learn different variants of the DMS task using the same ME and MS signals, and to flexibly adjust the decision criteria when the stimulus statistics are changed.

In the model, we do not assign the working memory, comparison, and decision-making operations to specific brain areas. The local cortical circuits for these three basic operations may be located within a single brain area, or be distributed across several areas. For example, subpopulations of neurons in the prefrontal cortex exhibit activities consistent with all operations involved in the DMS task: sample-selective delay activity, ME/MS comparison signals, and match/nonmatch decision signals (Miller et al., 1996; Freedman et al., 2002). However, ME and MS neural signals have also been observed in the parietal areas 7a (Rawley and Constantinidis, 2010), LIP and MIP (Swaminathan et al., 2010), in the inferior temporal cortex (Miller and Desimone, 1994; Freedman et al., 2003), and in the area V4 (Kosai et al., 2010). These areas differ in the magnitude, latency, and the proportion of neurons carrying each type of signal. This suggests that they are playing distinct or complementary roles in the match/nonmatch decision making, but which area is the source of comparison and decision signals remains to be elucidated in the future.

Active and passive comparison mechanisms

We first consider the dynamics of the comparison network (Fig. 2). The top–down input modulates neural activities without disrupting selectivity for motion direction. Neurons respond to their preferred test stimuli, but the response is higher in the ME cells than in the MS cells if the sample was also the preferred stimulus (match), and vice versa if the sample was the anti-preferred stimulus (nonmatch) (Fig. 2*A–C*). The ME and MS effects are specific for behavioral matches (i.e., for stimuli that match the sample stored in the WM circuit), as demonstrated by the responses to repeated nonmatch in Figure 2, *A* and *C*. The model thus reproduces the salient neural activity patterns observed in behaving monkeys (compare Figs. 2*C*, 1*D*). Interestingly, the model makes a testable prediction that the ME cells exhibit sample-selective delay activity (Fig. 2*A*,*B*). The delay activity in the ME neurons is induced solely by the top–down input, since the comparison network is dominated by recurrent inhibition and cannot sustain persistent firing on its own.

If the sample-tuned modulation from the WM circuit is disrupted (e.g., if the sample stimulus does not trigger persistent firing or if the top–down connections are absent), the active mechanism is abolished and the passive repetition suppression prevails in all cells in the comparison network (Fig. 2*D–F*; compare with experimental data in Fig. 1*C*). The passive mechanism does not distinguish behaviorally relevant and irrelevant repetitions; hence responses to match and repeated nonmatch are equally suppressed.

The circuit mechanism of active enhancement and suppression is illustrated in Figure 3, *A* and *B*. In the nonmatch condition (Fig. 3*A*), the bottom–up and top–down inputs target different columns in the ME population. The neurons tuned to the test stimulus are effectively driven by the bottom–up and recurrent inputs only. In this case, the ME cells have lower activity than the MS cells, since the recurrent and bottom–up synapses are weaker in the ME cells. In the match condition (Fig. 3*B*), the bottom–up and top–down inputs converge to the neurons within the same column. In this case, the top–down input compensates for the weaker recurrent excitation in the ME cells as well as for the adaptation-induced reduction in their responsiveness. Consequently, the ME cells show higher activity than the MS cells.

The dynamics in the comparison network have to be strongly dominated by recurrent feedback inhibition to achieve that the response to match stimuli is lower in the MS cells than in the ME cells. Indeed, in the match condition, the total activity of the ME

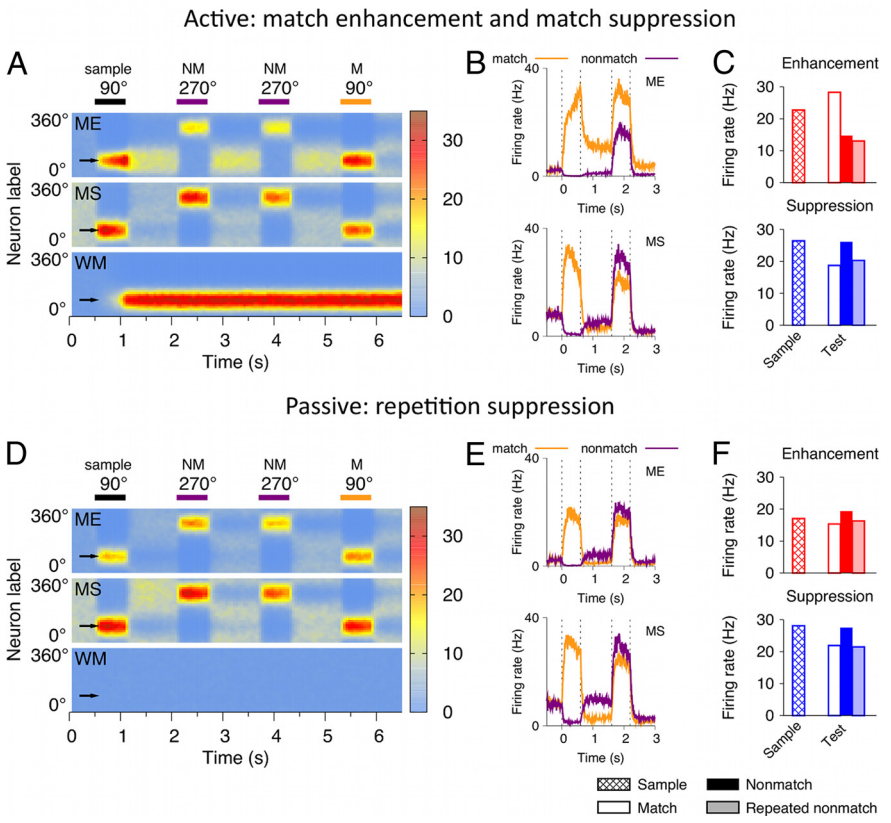


Figure 2. Active and passive memory mechanisms in the circuit model: active match enhancement and suppression (***A–C***), passive repetition suppression (***D–F***). ***A***, ***D***, Spatiotemporal activity pattern in the WM, ME, and MS populations in an ABBA task, where a sample (90°) is followed by two nonmatch test stimuli (270°) and then by the final match (90°). *x*-axis, Time; *y*-axis, neurons labeled by their preferred directions; firing rate is color-coded. ***A***, Comparison neurons respond to their preferred stimuli, but the activity is higher in the ME cells than in the MS cells for the match, and vice versa for the nonmatch stimuli. ***D***, If the activity in the WM circuit is disrupted, passive repetition suppression prevails in the comparison neurons. ***B***, ***E***, Firing rates of a neuron preferring the test stimulus on two trials: when the test appears as a match (orange line) and as a nonmatch (purple line). In the match condition, the sample is also the preferred stimulus for this neuron, and in the nonmatch condition the sample is the antipreferred stimulus. Note sample-selective persistent activity in the ME cell during the delay. ***C***, ***F***, Average responses to the preferred stimulus of the neuron appearing as a sample, match, nonmatch, and repeated nonmatch. These model results account for the single-neuron activities recorded from behaving monkeys in Figure 1, *C* and *D*.

and MS cells, and hence the recurrent excitation to the MS cells (that have stronger recurrent synapses), is comparable with that in the nonmatch condition. Nevertheless, in the match condition, the MS cells show lower activity than in the nonmatch condition. This is possible if the overall feedback inhibition is higher in the match than in the nonmatch condition. Since the feedback inhibition is approximately proportional to the summed activities of the ME and MS neurons, a signature of this network mechanism is that the total activity of the ME and MS cells is slightly higher in the match than in the nonmatch condition. In other words, the response of the ME cells in the nonmatch condition is lower than the response of the MS cells in the match condition (Figs. 2*C*, 3*C*). Our proposed mechanism of enhancement and suppression hence accounts for some subtle details of the experimental data shown in Figure 1 *D*. Notably, the overall activity in the comparison network is higher for match than for nonmatch stimuli despite the passive adaptation acting to reduce firing of the mostly active cells in the match condition.

When examining firing patterns in the comparison network, for all possible comparisons across different cell types and sample/match/nonmatch conditions, what matters the most is the difference in response of the ME and MS cells to the same test stimulus. This difference in firing of the ME and MS cells is what is used by the readout system to generate a categorical match versus nonmatch decision. The dynamical enhancement and suppression mechanisms in our model underlie this pattern of firing rate differences, which closely captures experimental data. In contrast, the exact responses to the sample in the ME and MS cells are not essential. In our model, neural responses to the sample are somewhat higher than the ME neural response to a nonmatch or MS response to a match test stimulus, which is attributable to the transient interplay of the rising activity in the WM circuit and of the building up adaptation current during the sample stimulus presentation as well as the enhanced global feedback inhibition afterward.

Sample-test similarity tuning in the ME and MS populations

So far, we considered only nonmatch stimuli that differed by 180° from the sample (i.e., the opposite direction of motion). It is interesting to see how the comparison network handles nonmatch stimuli with various degree of similarity to the sample. The directional difference between the sample and test determines the amount of overlap between the bottom–up and top–down inputs to the ME population (Fig. 3*A*). The larger this overlap is, the higher is the overall activity in the ME population. Accordingly, the response of the ME population is the highest in the match condition and gradually decreases with the directional difference, whereas the MS population exhibits the opposite pattern (Fig. 3*C*). In this way, neurons in the comparison network exhibit sigmoidal tuning to similarity between the sample and test, whereby activity of the ME cells increases and that of the MS cells decreases for more similar stimuli. Our model makes it explicit that similarity tuning is required to perform a DMS task and predicts that match enhancement and suppression effects are tuned to similarity in complementary ways. This predication can be tested experimentally.

Since a match/nonmatch decision is expected to rely on the differential signals from the ME and MS neurons, a key property of the network is the value of the directional difference at which the sample-test similarity tuning curves of the ME and MS cells cross. This value depends on the width of neural tuning in the WM and comparison networks, which is ~30°–50°, comparable with those observed in cortical neurons (Albright, 1984). Consequently, the two similarity tuning functions are coarse, and the crossing point is at ~70° (Fig. 3*C*, dashed vertical line). This raises two questions: Is the coarseness of the similarity tuning the main factor limiting the decision accuracy, and how can coarsely tuned neurons carry out fine discriminations? These questions are addressed in the following sections, in which we propose a downstream decision circuit that gener-

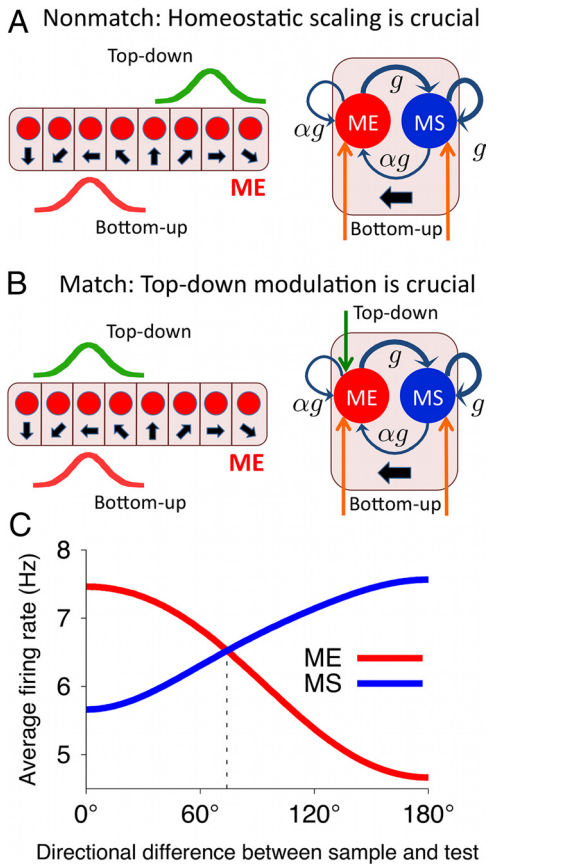


Figure 3. Circuit mechanism of match enhancement and suppression and neural tuning to the sample-test similarity. ***A***, ***B***, Left, Configuration of the top–down (green) and bottom–up (red) inputs to the ME population in the nonmatch (***A***) and match (***B***) conditions. Right, A column with the ME and MS neurons preferring the test stimulus is drawn. ***A***, Nonmatch condition, The MS neuron has higher activity because of stronger recurrent excitation (thick blue arrows). ***B***, Match condition, Top–down input compensates for weaker recurrent excitation, and the ME neuron has higher activity. ***C***, Similarity tuning. Average population firing rate for the ME (red line) and MS (blue line) neurons as a function of directional difference between the sample and test. The ME and MS populations are parametrically tuned to the sample-test similarity in complementary ways.

ates match/nonmatch choices based on activities of the ME and MS neurons.

Learning to compute the match/nonmatch decision

The decision circuit comprises two competing neural populations selective for the choices (e.g., match and nonmatch) (Fig. 1*E*) and exhibits winner-take-all dynamics. Across trials, the stochastic choice behavior in the decision circuit is characterized by a sigmoidal function, which represents how the probability of making a choice depends on the difference in synaptic input currents to the two competing neural populations (Soltani and Wang, 2006) (Fig. 4*B*). Since the ME and MS neurons are entangled within the comparison network and have the same cellular properties, it is reasonable to assume that the ME and MS neurons are all connected to both selective populations in the decision circuit. Specific connectivity profiles that differentially weight activities of the ME and MS neurons should emerge from experience-dependent learning (Fig. 4*A*).

We used a reward-dependent Hebbian learning rule similar to that in the previous work (Soltani and Wang, 2006, 2010; Fusi et al., 2007) (see Materials and Methods), but with the additional assumption that the synaptic potentiation/depression rate $q_0 \cdot q(r)$ is an increasing function of the presynaptic firing rate r (Fig. 4*C*). Since neurons in the decision circuit have binary (high or low) activities, for simplicity we reduce the dependence on the postsynaptic firing to a binary rule: only synapses onto the population with the high activity (i.e., for the winner that determines the choice) are updated. Synapses are potentiated in reward trials, and depressed in error trials.

Gradual dependence of $q(r)$ on the presynaptic firing is the key to learning the task. Consider a ME cell and a MS cell preferring the test stimulus, and consider their four connections to the match and nonmatch populations, c^{M}_{ME}, c^{NM}_{ME}, c^{M}_{MS}, c^{NM}_{MS} (Fig. 4*A*). If the test stimulus is a match, then the firing rate and hence the amount of potentiation/depression is slightly higher for the ME cell (Fig. 4*C*). The match choice is rewarded in this condition and induces potentiation in both cells, but synapses from the ME cell are potentiated more than those from the MS cell (Fig. 4*C*), leading to $c^{M}_{ME} > c^{M}_{MS}$. The nonmatch choice is not rewarded in this condition, and synapses from the ME cell are depressed more than the synapses from the MS cell, leading to $c^{NM}_{ME} < c^{NM}_{MS}$. The similar argument applies to the case of a nonmatching test. In this way, learning eventually gives rise to a synaptic connectivity profile such that the ME and MS neurons preferentially target the match and nonmatch populations, respectively (Fig. 4*A*,*D*).

If learning is performed with randomized direction of the sample stimulus, all motion directions are presented equally often during the test. As a consequence, the steady-state synaptic strength for each comparison neuron is independent of its preferred motion direction. That is, four values, c^{M}_{ME}, c^{NM}_{ME}, c^{M}_{MS}, c^{NM}_{MS}, fully characterize the steady state of the learning process (Fig. 4*D*). The steady-state values of synaptic strengths can be calculated analytically (see Materials and Methods), which in turn allow us to calculate the psychometric function of the network (Fig. 4*E*,*F*). The steady-state prediction is the upper bound on the behavioral performance. Ongoing learning in the network produces time-varying fluctuations of synaptic strengths around their steady-state values, which results in slightly lowered performance. The magnitude of these fluctuations increases with the maximal learning rate q_0. There is therefore a trade-off between faster learning and higher accuracy (Fig. 4*E*,*F*). For sufficiently low q_0, the performance approaches the steady-state level.

What determines the behavioral performance

Behavioral performance in our model is jointly determined by three factors: firing rates of neurons in the comparison circuit, sensitivity of the decision circuit, and the profile of synaptic connections between the comparison and decision networks. To discern contributions from each of these three factors, we computed the performance of the model, allowing one of them to vary while holding the remaining two factors fixed (Fig. 5). It is instructive to perform this analysis using linear similarity tuning in the ME and MS populations as well as linear dependence of the learning rate $q(r)$ on the firing rate, as we have assumed for the results in Figure 5. Linear similarity tuning allows us to determine and parametrically vary the sharpness of tuning through just a single parameter, the tuning slope α. Moreover, the slope α is the same for all directional differences and the accuracy at small directional differences is not constrained by the nonlinear saturation as it is the case for sigmoidal tuning.

First, consider how the performance of the model depends on the synaptic connectivity profile, with the parameters of the comparison and decision networks fixed. In Figure 5*B*, we plot the overall per-

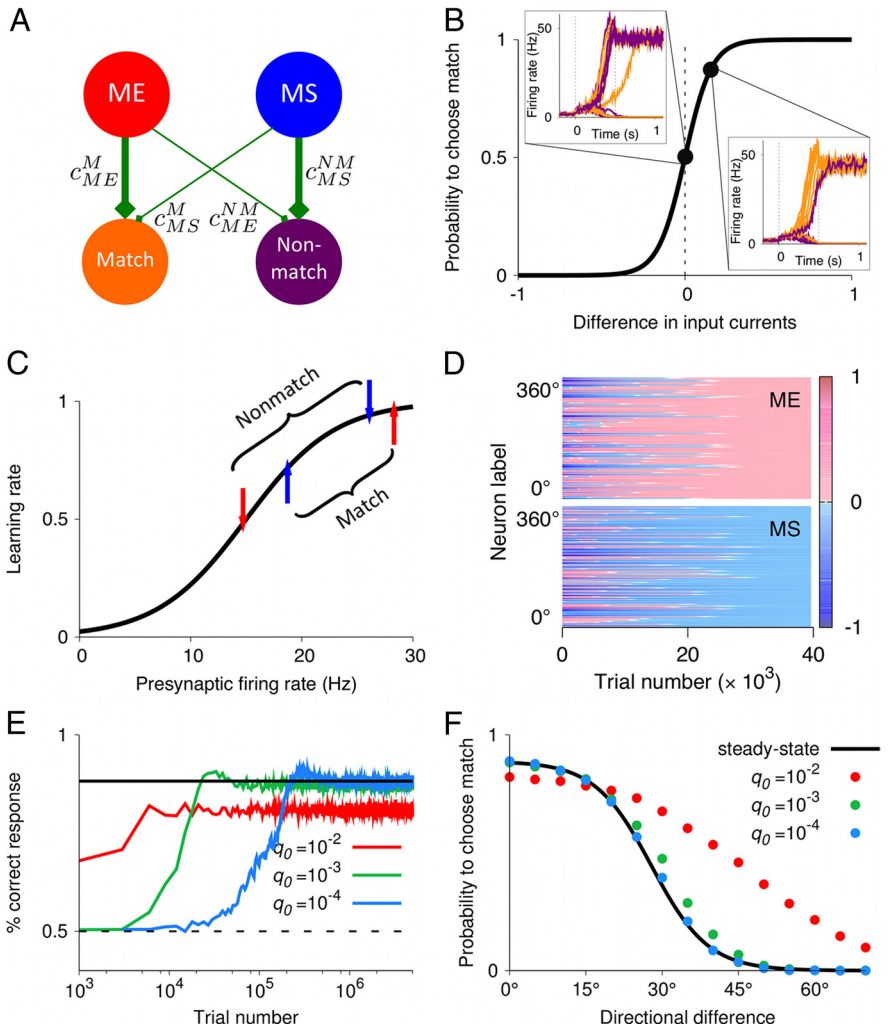


Figure 4. Learning the DMS task through reward-dependent Hebbian plasticity. ***A***, Schematic of synaptic connections between the comparison (ME and MS) and the decision (match and nonmatch) populations. Through synaptic plasticity, a connectivity profile emerges such that the ME and MS neurons preferentially target match and nonmatch populations, respectively (i.e., $\Delta c_{\mathrm{ME}} = c^{\mathrm{M}}_{\mathrm{ME}} - c^{\mathrm{NM}}_{\mathrm{ME}} > 0$ and $\Delta c_{\mathrm{MS}} = c^{\mathrm{M}}_{\mathrm{MS}} - c^{\mathrm{NM}}_{\mathrm{MS}} < 0$). ***B***, In the decision circuit, the trial-averaged performance is captured by the sigmoidal dependence of probability to choose match on the difference in synaptic input currents to the match and nonmatch populations, $\Delta I = g\Sigma_i[\Delta c_{\mathrm{ME}} r_{\mathrm{ME}} + \Delta c_{\mathrm{MS}} r_{\mathrm{MS}}]$. Firing rates of the match (orange) and nonmatch (purple) populations in 10 simulated trials are shown in two cases: for $\Delta I = 0$ when match and nonmatch are chosen equally often; for $\Delta I > 0$ when match is chosen more frequently than nonmatch. ***C***, Learning rate is a monotonically increasing function of the presynaptic firing rate. The arrows indicate the firing rates of a ME (red) and MS (blue) neuron in response to their preferred stimulus appearing as match and as nonmatch (0° and 180° directional difference, respectively). ***D***, Spatiotemporal dynamics of the synaptic strengths. Differences of the synaptic strengths Δc_{ME} and Δc_{MS} are color coded for all comparison neurons. *x*-axis, Trial number; *y*-axis, presynaptic neurons labeled by their preferred directions. ***E***, In the learning process, the fraction of correctly performed trials increases faster for higher learning rates q_0. Solid black line, Steady-state performance; dashed line, chance level. ***F***, Psychometric function obtained from the steady-state calculations (black line) and from simulations with different q_0 (colored circles). The performance approaches the steady-state level for sufficiently low q_0. Stimulus statistics is the same as in Figure 7 for $p_0 = 0.5$.

formance of the model (percentage correct responses) as a function of differences in synaptic strengths $\Delta c_{\mathrm{ME}} = c^{\mathrm{M}}_{\mathrm{ME}} - c^{\mathrm{NM}}_{\mathrm{ME}}$ and $\Delta c_{\mathrm{MS}} = c^{\mathrm{M}}_{\mathrm{MS}} - c^{\mathrm{NM}}_{\mathrm{MS}}$. The synaptic strengths in Figure 5*B* are not adjusted by learning, we rather ask how well does the model perform for given values of synaptic strengths. Note that $-1 \le \Delta c \le 1$, since the synapses are bounded $0 \le c \le 1$.

Probability of choices in the decision circuit depends on the difference in input currents to the match and nonmatch selective populations, $\Delta I = g[\Delta c_{\mathrm{ME}} r_{\mathrm{ME}} + \Delta c_{\mathrm{MS}} r_{\mathrm{MS}}]$. If both Δc_{ME} and Δc_{MS} have the same sign, which means that both ME and MS cells are more strongly connected to the same pool in the decision circuit, then ΔI has the same sign for all directional differences. In this case, the model always generates the same response and the performance is at chance level (Fig. 5*B*, green area). If $\Delta c_{\mathrm{ME}} < 0$ and $\Delta c_{\mathrm{MS}} > 0$, then the match response is more probable when the activity in the MS population is higher (i.e., for large directional differences) and less probable when the activity in the ME population is higher (i.e., for small directional differences). In this case, the performance is worse than chance (Fig. 5*B*, blue area). Finally, the region where $\Delta c_{\mathrm{ME}} > 0$ and $\Delta c_{\mathrm{MS}} < 0$ corresponds to the ME and MS cells being more strongly connected to the match and nonmatch populations, respectively. Here, the match response is more (less) probable for small (large) directional differences and the performance is higher than chance (Fig. 5*B*, yellow-to-red area).

Let us now see how within this region, where $\Delta c_{\mathrm{ME}} > 0$ and $\Delta c_{\mathrm{MS}} < 0$, the performance and the psychometric threshold of the model depend on the relative magnitudes of synaptic strengths, $\lambda = |\Delta c_{\mathrm{MS}}/\Delta c_{\mathrm{ME}}|$. In this region, the difference in synaptic currents can be rewritten as $\Delta I = |\Delta I_{\mathrm{ME}}| - |\Delta I_{\mathrm{MS}}|$, where $|\Delta I_{\mathrm{ME}}| = g|\Delta c_{\mathrm{ME}}|r_{\mathrm{ME}}$ and $|\Delta I_{\mathrm{MS}}| = g|\Delta c_{\mathrm{MS}}|r_{\mathrm{MS}}$. The dependence of these two contributions on the directional difference is obtained just by multiplying the similarity tuning curves of the ME and MS neurons by their respective $|\Delta c|$ values (Fig. 5*C*). The directional difference at which $|\Delta I_{\mathrm{ME}}|(\theta_i)$ and $|\Delta I_{\mathrm{MS}}|(\theta_i)$ curves cross corresponds to $\Delta I = 0$ [i.e., to $P(\mathrm{match}) = P(\mathrm{nonmatch}) = 0.5$] and is referred to as the point of subjective indifference (PSI). Let us see how PSI, and consequently the psychometric threshold, depend on the parameter λ. For $\lambda = 1$ (i.e., $|\Delta c_{\mathrm{MS}}| = |\Delta c_{\mathrm{ME}}|$), the two curves, $|\Delta I_{\mathrm{ME}}|(\theta_i)$ and $|\Delta I_{\mathrm{MS}}|(\theta_i)$, cross exactly at the same directional difference where $r_{\mathrm{ME}}(\theta_i)$ and $r_{\mathrm{MS}}(\theta_i)$ curves cross (Fig. 5*C*, orange lines). Since the similarity tuning in the ME and MS neurons is coarse, the PSI and the psychometric threshold are large (~90°) in this case. For $\lambda < 1$ (i.e., $|\Delta c_{\mathrm{MS}}| < |\Delta c_{\mathrm{ME}}|$), the crossing point of $|\Delta I_{\mathrm{ME}}|(\theta_i)$ and $|\Delta I_{\mathrm{MS}}|(\theta_i)$ shifts to even larger directional differences (Fig. 5*C*, blue line). Hence the PSI and the psychometric threshold increase, which is reflected in lower overall performance (Fig. 5*B*, yellow off-diagonal area). In contrast, for $\lambda > 1$ (i.e., $|\Delta c_{\mathrm{MS}}| > |\Delta c_{\mathrm{ME}}|$), the crossing point of $|\Delta I_{\mathrm{ME}}|(\theta_i)$ and $|\Delta I_{\mathrm{MS}}|(\theta_i)$ shifts to smaller directional differences (Fig. 5*C*, green line). The PSI and the psychometric threshold decrease and the overall performance increases (Fig. 5*B*, dark-red off-diagonal area) until the imbalance between $|\Delta c_{\mathrm{ME}}|$ and $|\Delta c_{\mathrm{MS}}|$ reaches the value where $|\Delta I_{\mathrm{ME}}| < |\Delta I_{\mathrm{MS}}|$ for all θ_i and the performance quickly drops to the chance level [the drop-off happens within the range of ΔI values in which

the choices in the decision network are stochastic (Fig. 5*B*, right panel)]. The performance drops off sharply because of the discontinuity in the correct response: 0° is the match, but any nonzero directional difference is a nonmatch. As long as the curves $|\Delta I_{ME}|(\theta_i)$ and $|\Delta I_{MS}|(\theta_i)$ cross just between 0° and the smallest nonmatch directional difference ψ_1 (which is 5° in Fig. 5), the performance is the best possible, but a small change in the synaptic strengths resulting in $|\Delta I_{ME}|(0°) < |\Delta I_{MS}|(0°)$ will cause the network to respond "nonmatch" to 0° directional difference and hence the chance level performance. Note that the reward-dependent learning naturally adjusts synaptic strengths (Fig. 5*B*, white star) and drives the network as close as possible to the best performance, but far enough from the drop-off boundary so that fluctuations of synaptic strengths do not result in the chance level performance.

Another overall trend is that the performance slightly improves for larger values of $|\Delta c|$. This is because larger Δc result in larger absolute values of ΔI, for which the choices of the decision network are less stochastic (Fig. 4*B*). For the parameters as in Figure 5*B*, the performance of ~100% correct can be achieved with large enough $|\Delta c|$. How well does our learning rule perform compared with what is optimally possible? The steady-state values of Δc resulting from the learning rule (Fig. 5*B*, white star) correspond to 95% correct performance, which is slightly less than optimally possible. This is because the absolute values of learned Δc are small. These values reflect the difference in the average firing rate of a cell on rewarded match and nonmatch trials, and since the similarity tuning is smooth, the steady-state Δc are small.

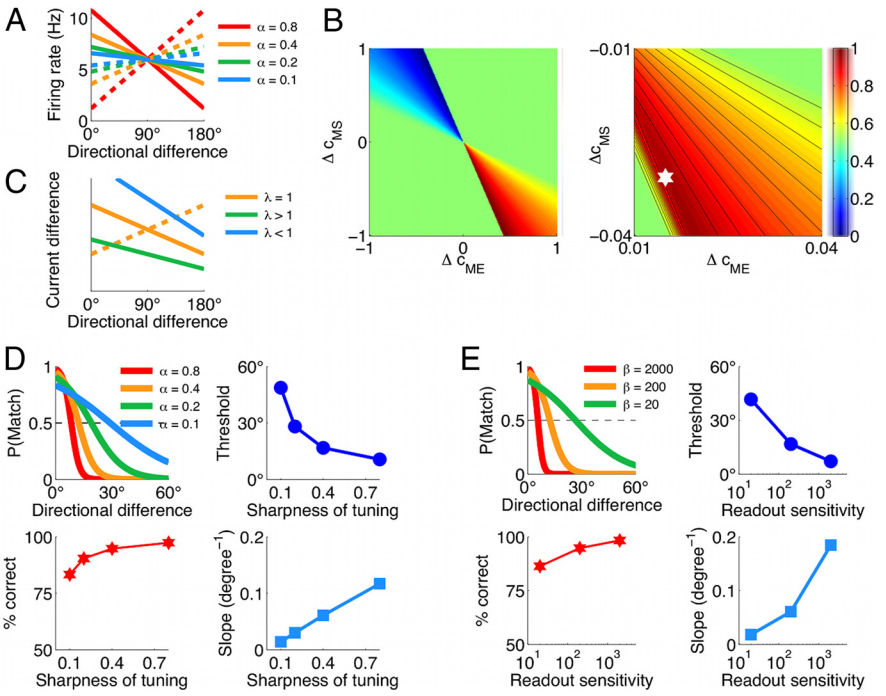


Figure 5. The behavioral performance of the model is jointly determined by the firing rates of the ME and MS neurons, sensitivity of the decision circuit, and the profile of synaptic connections between the comparison and decision circuits. For this simplified analysis, we assumed linear similarity tuning in the ME and MS populations as well as linear dependence of the learning rate $q(r)$ on the firing rate. Specifically, we used the functions $f_{ME,MS}(x) = \pm\alpha x + 0.5(1 \mp \alpha)$, where the upper and lower signs refer to the ME and MS populations, respectively. For different directional differences θ_i, the firing rates followed: $r_{ME,MS}(\theta_i/180°) = 12\ \text{Hz} \cdot f_{ME,MS}(x)$, and the learning rates were just $q_{ME,MS}(\theta_i/180°) = f_{ME,MS}(x)$. ***A***, The parameter α determines the sharpness of similarity tuning in the ME (solid lines) and MS (dashed lines) populations, whereby larger α corresponds to larger difference between the activities of the ME and MS populations. ***B***, The overall performance of the model (fraction of correct responses) color coded as a function of synaptic differences Δc_{ME} and Δc_{MS}. Right panel, Zoom into the region of small Δc. The white star indicates the steady-state Δc obtained through learning. $\alpha = 0.4$ and $\beta = 200$ nA^{-1} are fixed. ***C***, Two contributions to the difference in postsynaptic currents, $|\Delta I_{ME}|$ (solid lines) and $|\Delta I_{MS}|$ (dashed line) for different values of $\lambda = |\Delta c_{MS}/\Delta c_{ME}|$. The crossing point of these two curves and hence the psychometric threshold shift to larger directional differences for $\lambda < 1$, and to smaller directional differences for $\lambda > 1$. $\alpha = 0.4$. ***D***, Dependence of the psychometric function on the sharpness of similarity tuning in the comparison network. Sharper tuning (corresponds to larger values of α) results in lower psychometric threshold, larger slope of the psychometric function, and better overall performance. $\beta = 200$ nA^{-1} is fixed. ***E***, Dependence of the psychometric function on the sensitivity of the decision network. Higher sensitivity (corresponds to larger values of β) results in lower psychometric threshold, larger slope of the psychometric function, and better overall performance. $\alpha = 0.4$ is fixed. The synaptic strengths in ***D*** and ***E*** are adjusted through learning. Stimulus statistics is the same as in Figure 7 for $p_0 = 0.5$.

The dependence of the behavioral performance of the model on sharpness of the similarity tuning (parameter α) and on sensitivity of the decision circuit (parameter β) is presented in Figure 5, *D* and *E*, respectively. Here, synaptic strengths are adjusted through learning using linear ME and MS tuning curves (Fig. 5*A*). Shallower similarity tuning as well more stochastic decision circuit have similar effect on the behavior, producing decrease in the overall performance, increase in the psychometric threshold, and decrease in the slope of the psychometric function.

Degradation of performance with memory delay

In working-memory tasks, performance accuracy is known to decay with the duration of the memory delay (Pasternak and Greenlee, 2005). We propose that the main cause of worsened performance is degradation of the sample memory because of fluctuating neural dynamics in the WM circuit. After the sample stimulus is withdrawn, the WM circuit maintains its memory by reverberating activity. However, random fluctuations in the WM circuit can move elevated activity from one group of neurons to another, causing random drifts of the remembered sample (Fig. 6*B*). The variance of the sample memory grows linearly with time, consistent with a diffusion process (Camperi and Wang, 1998; Compte et al., 2000; Chow and Coombes, 2006; Carter and Wang, 2007) (Fig. 6*A*). Although a persistent activity pattern can be maintained for many seconds, the correlation between its peak (remembered sample) and the actual sample direction decays with time. Test stimuli are therefore compared with a corrupted memory of the sample, which leads to poorer performance (Fig. 6*C*,*D*). The model predicts that the psychometric threshold increases with the delay, in part because of a decrease in the slope of the psychometric function (i.e., decrease in sensitivity). The decay of relative discrimination with the memory delay (Fig. 6*D*) provides an explanation for several similar experimental observations (Pasternak and Greenlee, 2005).

Combining sensory evidence with priors by plastic synapses

In our model, synaptic modifications depend on the firing rates of neurons and the reward signal. Different statistics of stimuli used in the learning process entail a change in the statistics of firing rates. The ensuing plasticity could lead to synaptic strengths that adapt to the sensory environment and so optimize the network performance. Note that adapting to different stimulus statistics and task/reward rules does not require any change in

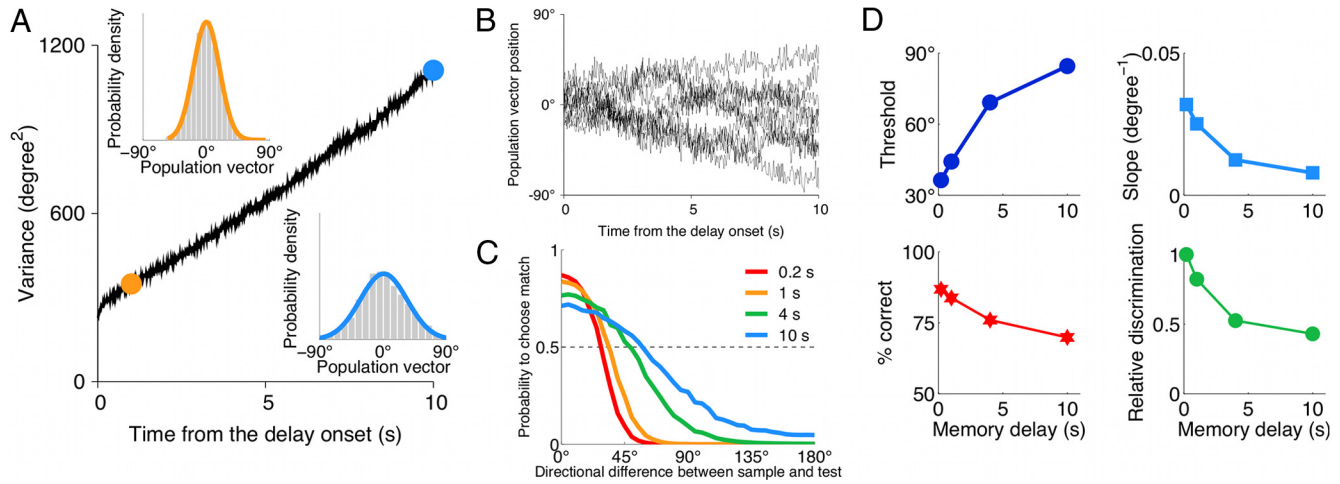


Figure 6. Degradation of performance in the DMS task with memory delay. ***A***, Memory of the sample is encoded by the peak location of the bell-shaped persistent activity pattern in the WM circuit (see Materials and Methods). Variance of the remembered sample growths linearly with time, consistent with a diffusion process. The insets show the probability density for the remembered sample after 1 s (orange) and 10 s (blue) delays (gray histogram, simulations; solid color line, Gaussian fit). ***B***, Example traces for the peak location of the persistent activity pattern in the WM circuit, which represents the sample memory during the delay. ***C***, Psychometric function in the DMS task for different durations of the memory delay. ***D***, Psychometric threshold increases and the slope of the psychometric function decreases for longer delays. The overall performance decreases for longer delays but remains at relatively high level for all delays. Relative discrimination (ratio of the threshold at 0.2 s delay to the threshold at longer delays) decreases with the delay duration, which accounts for the psychophysical observations with monkeys (Pasternak and Greenlee, 2005). Stimulus statistics is the same as in Figure 7 for $p_0 = 0.5$.

the model architecture or in the response properties of the ME and MS neurons. The same ME and MS neural signals can be used differently by the decision network because of flexible readout adjusted by reward-dependent plasticity.

Consider the impact of varying the prior probability p_0 that a test stimulus is match (Fig. 7). Evidently, changing the prior does not affect performance for easily discernible nonmatches with large directional differences (Fig. 7*B*). However, if the sample and test are very similar, then a nonmatch is difficult to be discriminated from the match. Indeed, the test–sample similarity (as well as the activity in the ME and MS pools) (Fig. 3*C*) changes smoothly with their directional difference, whereas the correct response exhibits a discontinuity: 0° is the match, but any nonzero (within given tolerance) directional difference is a nonmatch. Hence there is a trade-off: higher probability to correctly identify the match implies more errors on the nonmatches similar to the sample. To optimize performance, the behavior should be biased toward correct responses on the conditions (match or nonmatch) that are encountered more frequently.

Our plasticity rule naturally implements this trade-off (Fig. 7*B*). This is because synaptic modifications for a given stimulus contribute to the cumulative synaptic strength across trials in proportion to the frequency of its occurrence (Soltani and Wang, 2010). In this way, synaptic strengths encode priors (see Materials and Methods), which biases the behavior toward higher performance on stimuli that are more frequently encountered (Fig. 7*B*). The model makes a testable behavioral prediction that the psychometric threshold increases with the prior probability of the matching test (Fig. 7*D*), which is consistent with human psychophysics data (Vickers, 1979).

To compare with our neural circuit model, we computed performance of an ideal Bayesian observer (Fig. 7*C*) (see Materials and Methods). The network model and the ideal observer exhibit similar trends in how the psychometric function depends on the prior. The psychometric threshold (Fig. 7*D*), the probability to correctly identify match (Fig. 7*F*), and the slope of the psychometric function (Fig. 7*E*) increase for larger match prior p_0. Although changes in the psychometric function of the network model differ quantitatively from the ideal observer, their overall performance is virtually the same (Fig. 7*G*). For a low match prior p_0, this is because of the aforementioned trade-off: the ideal observer identifies match stimuli more accurately than the network model, but at the same time it produces more errors on nonmatch stimuli that are similar to the sample. For a high match prior p_0, the ideal observer performs better than the network model on nonmatch stimuli with intermediate directional differences ~30°–50°. However, because these stimuli occur very rarely when p_0 is high, there is no improvement in the overall performance. Therefore, we conclude that our biologically plausible mechanism achieves the same performance level as an ideal Bayesian observer.

Range of sample test similarities affects performance

Variations of the range of sample test similarities affects behavioral performance by implicitly changing priors for nonmatch stimuli that are similar to the sample. Consider a situation when the prior for a matching test is fixed at 0.5, but nonmatch similarity is varied by changing the range of directional differences used in the training (Fig. 8*A*). Nonmatches similar to the sample (5°–20°) appear less frequently when the distribution of directional differences is broader (Fig. 8*C*, gray bars). Since the synapses compute priors for all stimuli, the behavior again reflects the trade-off involved in discrimination of the match from very similar nonmatches (Fig. 8*B*,*C*). For a narrower range of directional differences, the accuracy of correctly identifying match is sacrificed for better performance on very similar nonmatches reflecting the increase in the prior probability for the latter (Fig. 8*C*). However, narrowing the range of directional differences also makes the task more difficult. The model predicts that the overall performance deteriorates with a decreased range of directional differences and eventually becomes just slightly above the chance level (Fig. 8*D*).

Adjusting the readout scheme to the task demands

The psychometric thresholds in Figures 7 and 8 are ~30°–60°, which agrees with the thresholds reported in monkey DMS par-

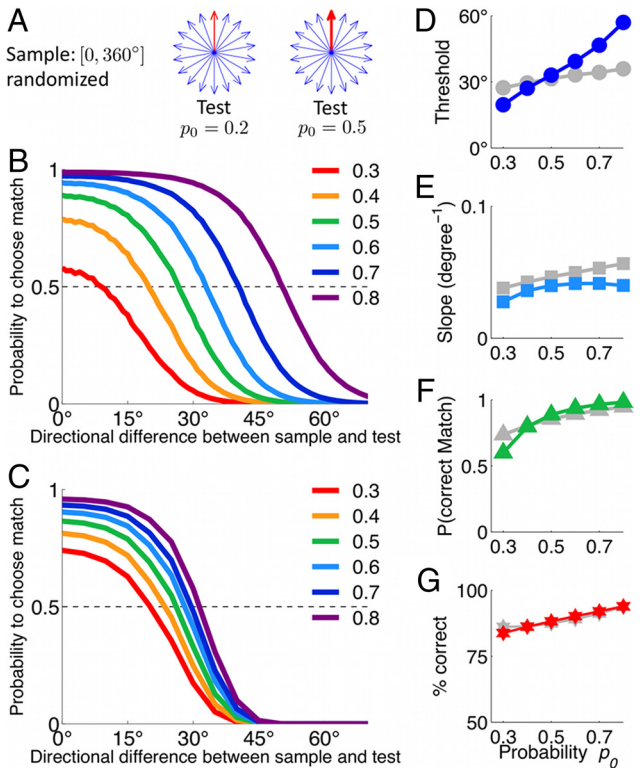


Figure 7. Plastic synapses encode priors for match and nonmatch and act to optimize performance. ***A***, Schematic of the stimulus statistics in the DMS task with different priors for match. Sample motion direction is drawn from a uniform distribution on [0°, 360°]. Match (red arrow) corresponds to zero directional difference. Nonmatches (blue arrows) differ from the sample by $\Delta\theta = \{\pm 5°, \pm 10° \ldots \pm 180°\}$, which are all equally probable. Note that the smallest nonmatch directional difference is ±5°, which sets the tolerance level. Match and nonmatch trials are randomly interleaved. Prior probability for a match trial is p_0 (indicated by the thickness of the red arrow). ***B***, Performance of the network model for different match priors p_0 (colored lines labeled by p_0 values). ***C***, Performance of the ideal Bayesian observer for different match priors p_0. In both cases (***B***, ***C***), the psychometric function changes toward higher probability to choose match as p_0 increases, which reflects the trade-off involved in fine discrimination between the match and nonmatch stimuli that are similar to the sample. ***D–G***, Psychometric threshold (***D***), slope of the psychometric function (***E***), probability to correctly identify match (***F***), and the overall performance (***G***) for the network model (colored symbols) and for the ideal Bayesian observer (gray symbols) as functions of the match prior p_0. Although changes in the psychometric function of the network model differ quantitatively from the Bayesian strategy, the overall performance of the network is virtually the same as for the ideal Bayesian observer.

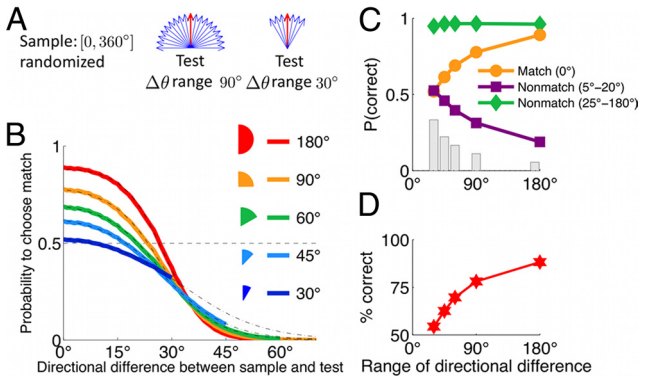


Figure 8. Range of sample test similarities affects performance on the DMS task. ***A***, Schematic of stimulus statistics with different ranges of sample-test similarities. Nonmatches differ from the sample by $\Delta\theta = \{\pm 5°, \pm 10° \ldots \pm \psi\}$, which are all equally probable, and ψ is the range of directional differences. Prior probability for a match trial is fixed at $p_0 = 0.5$. ***B***, As the range ψ decreases, the number of erroneous match decisions for small $\Delta\theta \neq 0°$ decreases, but the number of correct match decisions for $\Delta\theta = 0°$ also decreases. ***C***, Probabilities to correctly identify a match ($\Delta\theta = 0°$) and a nonmatch that is similar ($|\Delta\theta| = 5°–20°$) and dissimilar ($|\Delta\theta| = 25°–180°$) to the sample are plotted for five ψ values. The probability to correctly identify dissimilar, easily discernible nonmatch (green diamond) is always high. As the range ψ decreases, the probability to correctly identify very similar nonmatch (purple square) increases along with its prior probability (gray bar), whereas the probability to correctly identify match (orange circle) decreases. ***D***, Overall performance decreases as the range of directional differences becomes very narrow.

adigms (Zaksas and Pasternak, 2006) but is substantially larger than the thresholds of ~1°–2° reported in human and monkey fine discrimination paradigms (Hol and Treue, 2001; Purushothaman and Bradley, 2005). In fine motion discrimination, the sample typically has a fixed reference direction (e.g., upward), and the task is to judge whether a subtle deviation in the test direction is clockwise or counterclockwise relative to this reference (Purushothaman and Bradley, 2005). It has been proposed that not all neurons equally contribute to such fine discrimination decisions but that the neurons most sensitive to small changes in the relevant feature have the highest impact (Hol and Treue, 2001; Purushothaman and Bradley, 2005; Jazayeri and Movshon, 2007; Law and Gold, 2009). For fine motion discrimination, the most sensitive are neurons tuned 40°–70° away from the reference direction, so that the reference direction is on the "flank" of the tuning curve, where its slope and hence the sensitivity of the neuron is the highest. Psychophysical (Hol and Treue, 2001; Jazayeri and Movshon, 2007) and neurophysiological (Purushothaman and Bradley, 2005) evidence supports the idea that the activity of these "flanking" neurons is weighted more strongly in fine perceptual decisions; however, the underlying biophysical mechanism is unknown.

Such a mechanism naturally emerges in our model through plasticity of synapses onto the decision circuit. Using our model, we simulated a fine discrimination task in which the sample direction (reference) is fixed (e.g., upward), and the two decision neuronal populations now read out "clockwise" (CW) versus "counterclockwise" (CCW) choices (Fig. 9). Since neurons tuned to the reference direction fire at similar rates for clockwise and counterclockwise stimuli, their connections to the CW and CCW populations have similar strengths after learning ($\Delta c \approx 0$) (Fig. 9). Hence these neurons have little impact on the decision despite their high firing rate. In contrast, neurons tuned 40–70° away from the reference exhibit the largest difference between responses to clockwise and counterclockwise stimuli. As a result, their connections are stronger to the population encoding the choice (CW or CCW) associated with the higher firing rate, and weaker to the other population (Fig. 9). These neurons have larger Δc and hence higher impact on the decision. The fine discrimination performance of the model agrees well with experiments and reproduces a psychometric threshold of ~1°–2° (Fig. 9*B*). The key is learning with a fixed reference direction, which generates a synaptic profile that selectively emphasizes activity of neurons tuned 40°–70° away relative to this fixed reference. This is consistent with observations that fine discrimination learning often does not transfer between motion directions (Ball and Sekuler, 1987). In contrast, synaptic strengths are independent of neuronal tuning if the sample direction is randomized (Figs. 7, 8). Therefore, the same model can be used to perform different tasks because of synaptic plasticity that implements switching between different readout schemes according to task demands.

In a motion fine-discrimination task, different schemes of decoding neural responses in the area MT were evaluated for their ability to produce the observed psychophysical performance (Purushotha-

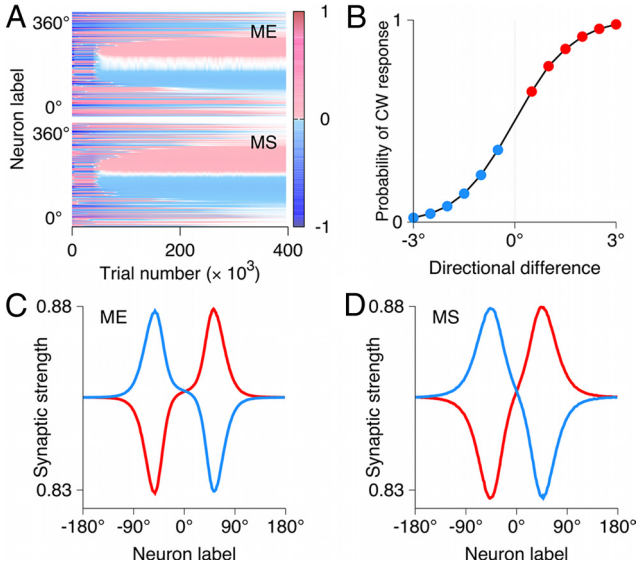


Figure 9. Synaptic plasticity adjusts the readout scheme according to task demands, illustrated by simulations of a fine motion discrimination (Purushothaman and Bradley, 2005). Sample moves in the fixed reference direction. Test stimuli are inclined by $\Delta\theta = \{\pm 0.5°, \pm 1° \ldots \pm 3°\}$ relative to the reference direction. The task is to judge whether a test stimulus is inclined clockwise ($\Delta\theta > 0$) or counterclockwise ($\Delta\theta < 0$) relative to the reference. After learning, the choice-selective populations in the decision circuit encode clockwise/counterclockwise (instead of match/nonmatch) decisions and hence are labeled as CW and CCW. ***A***, Spatiotemporal dynamics of the synaptic strengths. Differences of the synaptic strengths $\Delta c = c^{CW} - c^{CCW}$ are color coded for comparison neurons with all preferred directions. *x*-axis, Trial number; *y*-axis, presynaptic neurons labeled by their preferred directions. Through learning, a connectivity profile emerges, such that neurons tuned clockwise and counterclockwise relative to the reference preferentially target the CW- and CCW-selective populations, respectively. ***B***, Psychometric function for the fine motion discrimination. Psychometric threshold is ~1°–2°. ***C***, ***D***, Strengths of synaptic connections to the CW-selective (red; c_{ME}^{CW} and c_{MS}^{CW}) and CCW-selective (blue; c_{ME}^{CCW} and c_{MS}^{CCW}) populations after learning. Activity of each neuron is gradually weighted in the decision process, whereby higher weights are assigned to the most sensitive neurons tuned 40°–70° away from the reference direction.

man and Bradley, 2005). Predictions of our model agree with the conclusion of this analysis: fine-discrimination thresholds of a few degrees can only be achieved by the readout schemes that assign higher weights to neurons tuned away from the reference direction, but not by broad equal-weight schemes (Purushothaman and Bradley, 2005). Moreover, our model demonstrates a simple and realistic neural circuit for such a readout.

Comparison with a one-pool model

The match/nonmatch decisions in our model are based on the differential activity of ME and MS populations tuned to similarity in complementary ways (Fig. 10*A*, two-pool comparison model). These two complementary populations have been observed in neurophysiological studies of behaving monkeys. However, one may wonder whether ME and MS neurons are redundant, and whether only one of these two populations might be sufficient to perform the match versus nonmatch computation. Indeed, in an alternative scenario (Carpenter and Grossberg, 1987, 2003), a single neural population performs a simple addition of a sensory test input and an input representing the sample stimulus, and a match or nonmatch decision is determined by whether the converging inputs exceed a threshold (Fig. 10*A*, one-pool addition model). We compared our two-pool model with an implementation of the one-pool addition model. The latter is similar to the former, except that the intermediate layer consists of a single class of neurons, which are all driven by sensory and WM inputs (see Materials and Methods, Alternative model). Larger overlap between the WM and sensory signals results in higher overall activity in the addition population. Hence the activity of the addition population in the one-pool model monotonically decreases with directional difference between the sample and test, resembling similarity tuning curve of the ME neurons in the two-pool model (Fig. 10*B*,*C*, solid black lines). A downstream system can then read out match/nonmatch decisions by detecting whether the overall activation in this single population exceeds a threshold value (see Material and Methods).

The performance of the one-pool model is not robust against fluctuations in the strength of sensory inputs. Consider a situation when the strength of the sensory input increases on a trial, for example because of change in the contrast of visual stimulus. Neurons in both models respond with higher firing rates to stronger stimuli (Fig. 10*B*,*C*). Since the decision readout in the one-pool model relies on the absolute value of the firing rate in a single neural population, stronger sensory input will produce a drop in behavioral performance and increase in the psychometric threshold (Fig. 10*D*). In the two-pool model, however, the readout is based on differential activity of the ME and MS populations and not on the absolute value of their firing rates. Firing rates of both ME and MS populations equally increase in response to stronger inputs, but the behavioral performance of the two-pool model remains almost unaffected by changes in the strength of sensory input (Fig. 10*D*). In the same vein, noise in the sensory input equally affects firing of the ME and MS neurons and hence does not strongly impact behavioral performance in the two-pool model, whereas performance of the one-pool model is sensitive to input noise.

It is worth noting that the architecture of the one-pool addition model is not substantially simpler than the two-pool comparison model: it also requires a WM module to store the sample stimulus, an intermediate neural layer, and a readout system for match/nonmatch decisions. Importantly, we found that the one-pool model is vulnerable to variations of the strength of sensory stimuli, whereas the performance of the two-pool model is very robust, suggesting functional advantages of the two-pool comparison mechanism. Furthermore, the readout system in the two-pool model allows for flexible mapping between the decision and motor response. Behavioral tasks may require to respond for match only, for nonmatch only, or to indicate match and nonmatch by different responses. In the two-pool comparison model, match and nonmatch decisions are encoded in activity of two complementary neural populations. This activity is sufficient to drive an arbitrary motor response. By contrast, in the one-pool addition model, the readout unit is only activated for match decisions, and nonmatch decisions are represented by the lack of activity. If response for nonmatch is required by the task, there is no neural activity to drive such a motor response, and it is problematic to justify how it can be generated without additional model assumptions.

Discussion

In this paper, we proposed a recurrent neural circuit model for match versus nonmatch pattern comparison that is capable of performing all the key computations for DMS tasks. Similarity between the sample and test stimuli is encoded by the magnitude of response modulations (ME and MS) in two subpopulations of neurons within the comparison network. The test–sample similarity tuning in these cells arises from interactions of bottom–up and top–down inputs and strong local feedback inhibition. Similar-

ity signals are then pooled through plastic synapses by a downstream decision circuit that generates categorical match or nonmatch decisions. Using the same ME and MS neural signals, learning enables the network to generate decisions flexibly depending on stimulus statistics and task/reward rules in different behavioral tasks.

Alternative models for match versus nonmatch computation

For the DMS task, the one-pool "addition model" (Fig. 10*A*, right panel) (Carpenter and Grossberg, 1987, 2003) is intuitively plausible, but physiological data from behaving animals suggest a different scenario involving two (ME and MS) neural populations (Miller et al., 1996; Freedman et al., 2003; Rawley and Constantinidis, 2010). Our results indicate that the two-population architecture provides a more robust and flexible way to compute match versus nonmatch decisions than the one-pool architecture. Other models for the DMS task have previously been proposed, which rely on comparison mechanisms other than similarity tuning of two complementary populations. One model (Sugase-Miyamoto et al., 2008) ascribes the comparison computation to neurons acting as matched filters: on each trial, they rapidly adjust the strengths of input synapses to match the magnitude of individual inputs from the sample stimulus. Inputs from the test stimuli are then multiplied by those synaptic strengths, thereby computing a measure of the similarity between the sample and test. The biophysical mechanism underlying such one-shot learning is unclear. Moreover, a hypothetical "learn" command has to be introduced to trigger plasticity for the sample but not for the test stimuli. In another model (Tagamets and Horwitz, 1998), match response is triggered whenever a coactivation of two units is detected: one transiently responding to all sensory stimuli and another showing rising activity during the delay period but silent during the sample period. This model fails on the ABBA type of task (it responds equally to match and repeated nonmatch) and predicts poor performance for brief delays such that the rising population has not reached high activity yet. This is in contrast with predictions of our model and with experimental data (Miller and Desimone, 1994; Pasternak and Greenlee, 2005).

Neural circuit models have also been proposed for the smaller versus larger comparison of two analog quantities (Machens et al., 2005; Miller and Wang, 2006); however, neither of these models can be adapted for the match versus nonmatch comparison. Match/nonmatch computation can be performed with arbitrary stimuli, for example with visual objects, which requires comparison of patterns rather than analog quantities. The circuit mechanism in our model realizes comparison of two activity patterns, elicited by the sample and test stimuli, and may be generalized to other types of stimuli and encoding schemes.

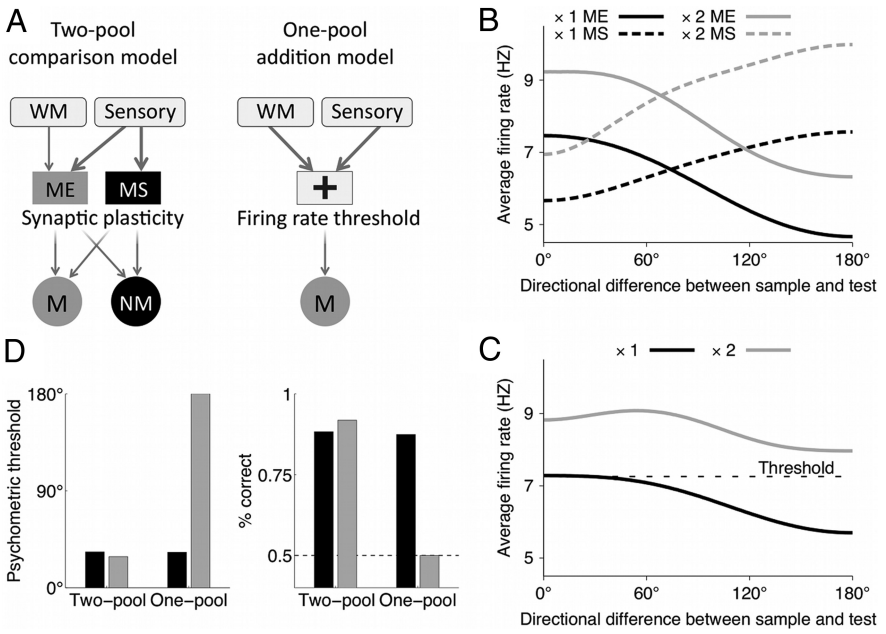


Figure 10. Behavioral performance in the two-pool comparison model, but not in the one-pool addition model, is robust to changes in the sensory input strength. ***A***, Schematics of the two-pool comparison model (simplified version of Fig. 1*E*) and of the one-pool addition model (for details, see Materials and Methods). ***B***, Average population firing rate for the ME (solid line) and MS (dashed line) neurons in the two-pool model as a function of directional difference between the sample and test. Black line, Control; gray line, doubled sensory input strength. The difference in the activity of ME and MS neurons is only slightly affected by the increase in the input strengths, whereas the firing rates in both populations increase significantly. ***C***, Average population firing rate for the addition population in the one-pool model as a function of directional difference between the sample and test. Black line, Control; gray line, doubled sensory input strength. The black dashed line indicates the firing rate threshold for match versus nonmatch decisions, obtained by fitting the parameters of the readout (Eq. 14) so as to match the psychometric functions for the one- and two-pool models in the control condition. ***D***, In the two-pool model, the psychometric threshold and overall performance remain almost the same for the control (black bar) and doubled (gray bar) input strength. In the one-pool model, the overall performance decreases and the psychometric threshold increases with the input strength. For the doubled input strengths (gray bar), performance drops to the chance level (dashed line), and the psychometric threshold (defined at 75% correct performance) cannot be determined; for comparison purpose, we plot the maximum possible threshold value, 180°.

Active and passive memory mechanisms

The active match enhancement mechanism in our model has several signatures that can be tested experimentally. First, the magnitude of enhancement and suppression encodes similarity between the sample and test. Second, neurons that show match enhancement also show sample-selective persistent activity during a mnemonic delay. This is consistent with the experimental observation that delay activity and match/nonmatch selectivity are present in mostly overlapping groups of neurons (Miller et al., 1996; Freedman et al., 2003) but needs to be verified more rigorously. Finally, the total neural activity (excitation and inhibition) is higher in the match than in the nonmatch condition. This suggests an explanation for the observations that functional magnetic resonance imaging signals in the relevant brain regions are larger for behavioral matches (Duncan et al., 2009), whereas neurophysiological data indicate that approximately equal proportions of cells show match enhancement and match suppression (Miller et al., 1996).

The passive repetition suppression is implemented in our model as the spike rate adaptation mediated by a long-lasting hyperpolarizing current. In single neurons, afterhyperpolarization effects lasting for multiple seconds can be mediated through a Na^+-activated K^+ current (Sanchez-Vives et al., 2000; Wang et al., 2003; Pulver and Griffith, 2010). An alternative or complementary mechanism may be synaptic depression (Chance et al., 1998). Long-term depression is involved in visual recognition

memory (Griffiths et al., 2008) producing suppressed responses to repeated stimuli. However, the biophysical mechanism underlying repetition suppression and recognition memory is still a matter of debate (Brown and Xiang, 1998).

In neurophysiological data as well as in our model, comparison neurons exhibit mixed selectivity (i.e., they are activated by a conjunction of the stimulus and match/nonmatch context). The mixed selectivity is essential for the match/nonmatch computation, which is an "exclusive or" (XOR) operation and belongs to the class of linearly nonseparable problems (Rigotti et al., 2010).

The active match enhancement in our model depends on the top–down modulation from the WM circuit. If this modulation is disrupted, passive repetition suppression prevails. We simulated this as a condition when the sample does not trigger persistent firing in the WM circuit, presumably because of the lack of attention. In neurophysiological data, repetition suppression dominates neural activity in animals trained on a simple feature matching task (Miller et al., 1991; Constantinidis and Steinmetz, 2001). When tested on a more complicated task, these animals initially respond to all perceptual matches regardless of their behavioral relevance (Miller and Desimone, 1994) (e.g., repeated nonmatch BB in the ABBA sequence). Extensive retraining is required to perform the task correctly, after which neural activity exhibits match enhancement (Miller and Desimone, 1994). Since persistent activity is present in the prefrontal cortex even during the standard DMS tasks, the retraining may be required to instruct the subjects about the rules of the new (ABBA) task, or/and involve rewiring of neural circuits, such as potentiation of top–down projections from the WM circuit. In the latter case, depending on the degree of this potentiation, repetition suppression can be still prevalent in cells showing persistent activity, or be supplemented by match enhancement when it is behaviorally required.

Computing decisions through plastic synapses

In our model, the comparison operation and match/nonmatch decision making are performed by separate neural networks that are connected through plastic synapses. This architecture allows the network to use the same ME and MS neural signals to perform different tasks and flexibly adjust decision criteria so as to optimize performance. In particular, when the prior probability of the matching test is increased, the model makes the testable prediction that the psychometric threshold increases with the prior.

A learning rule similar to ours (Soltani and Wang, 2010) was shown to be a linear approximation to a "Bayesian–Hebb learning rule" (Pfeiffer et al., 2010), which was designed specifically so that each synaptic weight converges to the log odds of receiving a reward when its presynaptic and postsynaptic neurons are active. This plasticity rule has been shown to successfully capture behaviors in probabilistic inference (Soltani and Wang, 2010), foraging (Soltani and Wang, 2006), and associative learning tasks (Fusi et al., 2007). In these previous works, the decision network was modeled in the same way as ours, but the Hebbian plasticity was implemented simply with a binary (all or none) dependence on the firing rate. By contrast, we propose that the learning rate is a graded increasing function of the presynaptic firing, which is critical for harnessing small differences in the neural signals of ME and MS neurons in the learning process. The activity-based weighting of neural responses enables the network to flexibly reconfigure the readout scheme according to task demands, for example so that activities of the flanking neurons are emphasized in a fine discrimination task.

It has been proposed that different readout schemes can be realized in a system in which neurons are tuned to the likelihoods of sensory stimuli (Jazayeri and Movshon, 2006). A discrimination decision can be made by comparing activities in a pair of neurons that encode the likelihoods of two directions to be discriminated. In such a model, different pairs of neurons are used for the coarse (e.g., 0° vs 180°) and fine (e.g., 0° vs 12°) motion discrimination, but the mechanism for selection of a particular pair from the likelihood-tuned population is unspecified. In contrast, different readout schemes in our model emerge just from stimulus statistics without any additional assumptions.

In conclusion, our model has identified simple biophysical mechanisms that, working together, are sufficient to carry out comparison between top–down expectation and bottom–up sensory stimulus pattern, leading to flexible match versus nonmatch decisions. Working memory, comparison, and categorical choice computations are basic "building blocks" of cognition; therefore, we expect that the general computational principles presented in this paper can be extended to delayed nonmatch-to-sample tasks and other types of match/nonmatch processes in the brain.

References

Abbott LF, Chance FS (2005) Drivers and modulators from push-pull and balanced synaptic input. Prog Brain Res 149:147–155.

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. J Neurophysiol 52:1106–1130.

Ball K, Sekuler R (1987) Direction-specific improvement in motion discrimination. Vision Res 27:953–965.

Born RT, Bradley DC (2005) Structure and function of visual area MT. Annu Rev Neurosci 28:157–189.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765.

Brown MW, Xiang JZ (1998) Recognition memory: neuronal substrates of the judgement of prior occurrence. Prog Neurobiol 55:149–189.

Camperi M, Wang XJ (1998) A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. J Comput Neurosci 5:383–405.

Carpenter GA, Grossberg S (1987) A massively parallel architecture for a self-organizing neural pattern recognition machine. Comput Vision Graph Image Process 37:54–115.

Carpenter GA, Grossberg S (2003) Adaptive resonance theory. In: The handbook of brain theory and neural networks, Ed 2 (Arbib MA, ed), pp 87–90. Cambridge, MA: MIT.

Carter E, Wang XJ (2007) Cannabinoid-mediated disinhibition and working memory: dynamical interplay of multiple feedback mechanisms in a continuous attractor model of prefrontal cortex. Cereb Cortex 17:i16–26.

Chance FS, Nelson SB, Abbott LF (1998) Synaptic depression and the temporal response characteristics of V1 cells. J Neurosci 18:4785–4799.

Chow CC, Coombes S (2006) Existence and wandering of bumps in a spiking neural network model. SIAM J Appl Dyn Syst 5:552–574.

Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb Cortex 10:910–923.

Constantinidis C, Steinmetz MA (2001) Neuronal responses in area 7a to multiple stimulus displays: II. Responses are suppressed at the cued location. Cereb Cortex 11:592–597.

Dubner R, Zeki SM (1971) Response properties and receptive fields of cells in an anatomically defined region of superior temporal sulcus in monkey. Brain Res 35:528–532.

Duncan K, Curtis C, Davachi L (2009) Distinct memory signatures in the hippocampus: intentional states distinguish match and mismatch enhancement signals. J Neurosci 29:131–139.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2002) Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. J Neurophysiol 88:929–941.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of

primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci 23:5235–5246.
Fusi S (2002) Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. Biol Cybern 87:459–470.
Fusi S, Asaad WF, Miller EK, Wang XJ (2007) A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. Neuron 54:319–333.
Griffiths S, Scott H, Glover C, Bienemann A, Ghorbel MT, Uney J, Brown MW, Warburton EC, Bashir ZI (2008) Expression of long-term depression underlies visual recognition memory. Neuron 58:186–194.
Gutkin BS, Laing CR, Colby CL, Chow CC, Ermentrout GB (2001) Turning on and off with excitation: the role of spike-timing asynchrony and synchrony in sustained neural activity. J Comput Neurosci 11:121–134.
Hol K, Treue S (2001) Different populations of neurons contribute to the detection and discrimination of visual motion. Vision Res 41:685–689.
Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. Nat Neurosci 9:690–696.
Jazayeri M, Movshon JA (2007) Integration of sensory evidence in motion discrimination. J Vis 7:7.1–7.
Kosai Y, Bushnell B, Pasupathy A (2010) V4 neurons reflect behavioral choice on a shape matching task. Soc Neurosci Abstr 36:892.5.
Kumaran D, Maguire EA (2007) Which computational mechanisms operate in the hippocampus during novelty detection? Hippocampus 17:735–748.
Law CT, Gold JI (2009) Reinforcement learning can account for associative and perceptual learning on a visual-decision task. Nat Neurosci 12:655–663.
Lisman JE, Grace AA (2005) The hippocampal-VTA loop: controlling the entry of information into long-term memory. Neuron 46:703–713.
Machens CK, Romo R, Brody CD (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. Science 307:1121–1124.
Miller EK, Desimone R (1994) Parallel neuronal mechanisms for short-term memory. Science 263:520–522.
Miller EK, Li L, Desimone R (1991) A neural mechanism for working and recognition memory in inferior temporal cortex. Science 254:1377–1379.
Miller EK, Li L, Desimone R (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. J Neurosci 13:1460–1478.
Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16:5154–5167.
Miller P, Wang XJ (2006) Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli. Proc Natl Acad Sci U S A 103:201–206.
Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. Nat Rev Neurosci 6:97–107.
Pfeiffer M, Nessler B, Douglas RJ, Maass W (2010) Reward-modulated hebbian learning of decision making. Neural Comput 22:1399–1444.
Pulver SR, Griffith LC (2010) Spike integration and cellular memory in a rhythmic network from Na^+/K^+ pump current dynamics. Nat Neurosci 13:53–59.
Purushothaman G, Bradley DC (2005) Neural population code for fine perceptual decisions in area MT. Nat Neurosci 8:99–106.
Rawley JB, Constantinidis C (2010) Effects of task and coordinate frame of attention in area 7a of the primate posterior parietal cortex. J Vis 10:12.1–16.
Renart A, Song P, Wang XJ (2003) Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. Neuron 38:473–485.
Rigotti M, Rubin DB, Wang XJ, Fusi S (2010) Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. Front Comput Neurosci 4:24.
Sanchez-Vives MV, Nowak LG, McCormick DA (2000) Cellular mechanisms of long-lasting adaptation in visual cortical neurons *in vitro*. J Neurosci 20:4286–4299.
Soltani A, Wang XJ (2006) A biophysically based neural model of matching law behavior: melioration by stochastic synapses. J Neurosci 26:3731–3744.
Soltani A, Wang XJ (2010) Synaptic computation underlying probabilistic inference. Nat Neurosci 13:112–119.
Steinmetz MA, Constantinidis C (1995) Neurophysiological evidence for a role of posterior parietal cortex in redirecting visual attention. Cereb Cortex 5:448–456.
Sugase-Miyamoto Y, Liu Z, Wiener MC, Optican LM, Richmond BJ (2008) Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. PLoS Comput Biol 4:e1000073.
Swaminathan S, McClellan S, Freedman D (2010) Comparing the roles of the prefrontal, lateral intraparietal, and medial intraparietal cortices in category-based decision making. Soc Neurosci Abstr 36:279.2.
Tagamets MA, Horwitz B (1998) Integrating electrophysiological and anatomical experimental data to create a large-scale model that simulates a delayed match-to-sample human brain imaging study. Cereb Cortex 8:310–320.
Turk-Browne NB, Yi DJ, Leber AB, Chun MM (2007) Visual quality determines the direction of neural repetition effects. Cereb Cortex 17:425–433.
Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature 391:892–896.
Vickers D (1979) Decision processes in visual perception. New York: Academic Professional.
Wang XJ (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24:445–463.
Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36:955–968.
Wang XJ, Liu Y, Sanchez-Vives MV, McCormick DA (2003) Adaptation and temporal decorrelation by single neurons in the primary visual cortex. J Neurophysiol 89:3279–3293.
Wolpert DM, Miall RC (1996) Forward models for physiological motor control. Neural Netw 9:1265–1279.
Wong KF, Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. J Neurosci 26:1314–1328.
Zaksas D, Pasternak T (2006) Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. J Neurosci 26:11726–11742.

Coherent stochastic oscillations enhance signal detection in spiking neurons

Tatiana A. Engel,[1] Brian Helbig,[2] David F. Russell,[3] Lutz Schimansky-Geier,[4] and Alexander B. Neiman[2]
[1]*Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA*
[2]*Department of Physics and Astronomy, Ohio University, Athens, Ohio 45701, USA*
[3]*Department of Biological Sciences, Ohio University, Athens, Ohio 45701, USA*
[4]*Institut für Physik, Humboldt Universität Berlin, Newton Str. 15, 12489 Berlin, Germany*
(Received 10 January 2009; revised manuscript received 30 June 2009; published 18 August 2009)

We study the effect of noisy oscillatory input on the signal discrimination by spontaneously firing neurons. Using analytically tractable model, we contrast signal detection in two situations: (i) when the neuron is driven by coherent oscillations and (ii) when the coherence of oscillations is destroyed. Analytical calculations revealed a region in the parameter space of the model where oscillations act to reduce the variability of neuronal firing and to enhance the discriminability of weak signals. These analytical results are employed to unveil a possible role of coherent oscillations in peripheral electrosensory system of paddlefish in improvement of detection of weak stimuli. The proposed mechanism may be relevant to a wide range of phenomena involving coherently driven oscillators.

DOI: 10.1103/PhysRevE.80.021919 PACS number(s): 87.19.ln, 05.40.−a, 87.19.lo

I. INTRODUCTION

Rhythmic activity is observed at all levels of nervous systems [1]. Recently, neuronal oscillations have been shown to improve the stimulus discrimination, either by increasing the precision of spike timing [2] or by inducing binomial statistics of serially binned spike counts across a neuronal population [3]. Here, we explore a different mechanism of discrimination enhancement, whereby coherent stochastic oscillations driving a periodically spiking neuron induce long-lasting serial correlations of interspike intervals (ISIs) that suppress variability of spike counts. This situation is encountered, for example, by a central nervous system neuron receiving coherent oscillatory inputs in gamma band, or in "hair cell–afferent" sensory receptors for hearing, balance, and electrosense.

The significance of anticorrelated ISIs in shaping neuronal responses was experimentally demonstrated in electroreceptors of a weakly electric fish [4,5] where negative correlations of sequential ISIs originate from intrinsic properties of sensory neurons. Theoretical studies have shown that a negative feedback mechanism either in a network [6] or in a single neuron with a dynamic threshold [7,8] can induce anticorrelations of sequential ISIs, which reduce long-time variability and thereby lead to enhanced information transfer of low-frequency stimuli. In this scenario, the correlation structure is fixed by the negative feedback parameters.

The long-lasting ISI anticorrelations can also result from the interaction of unidirectionally coupled noisy oscillators, as was shown in peripheral electroreceptors of paddlefish [9]. In this system, noisy oscillations in sensory epithelia drive sensory neurons and act as an internal narrow-band noise that leads to reduced low-frequency fluctuations of neuronal firing [10,11]. The structure and extent of serial ISI correlations in this feed-forward scenario are set by the frequency and coherence of driving oscillations. This could potentially provide an active adjustment mechanism to achieve enhanced discrimination on behaviorally relevant time scales.

The structure of ISI correlations in the neuron driven by slow exponentially correlated noise was also explored by several analytical studies using integrate and fire models [12–14]. It was found that exponentially correlated noise leads to positive serial ISI correlations, which enhance variability of neuronal firing at long time scales. The Fano factor, a measure of spike-count variability on different time scales, shows a minimum, indicating an optimal time scale at which variability of spike counts is lowest and hence the detection of weak signals is optimal [12].

Here, we aim to analyze how the signal detection depends on the correlation structure, as well as to contrast the signal detection by a spiking neuron driven by either coherent stochastic oscillations or by fluctuations whose coherence is destroyed. In both latter cases, the neuron firing is constrained to have identical first-order ISI statistics, but it lacks any ISI correlations if the oscillation coherence is destroyed. Using an analytically tractable model we calculate the parameter region where coherent stochastic oscillations enhance signal detection. The analytical model predictions are then verified using the experimental data from paddlefish electroreceptors [9].

II. METHODS AND MODEL

A. Spike-count variability and discriminability

To assess the performance of a spontaneously firing neuron as a detector for weak signals s, we apply the ideal observer paradigm [15] to the neuron's output spike train $y(t)=\Sigma_i \delta(t-t_i)$, where $\{t_i\}$ is the sequence of spike times. The observable is the spike count $n(T)=\int_t^{t+T} y(t')dt'$ in a parameter-fixed time window T. Decision about the presence of a constant signal s is made on the basis of comparing spike count distributions $P_s(n,T)$ and $P_0(n,T)$ in two situations: when a neuron is stimulated or spontaneously firing, respectively. The separation of the two distributions is characterized by the discriminability measure

$$d' = 2[|\mu_s(T) - \mu_0(T)|]/[\sigma_s(T) + \sigma_0(T)], \quad (1)$$

where μ and σ are the mean and the standard deviation of the spike count, respectively. The variability of the spike count is assessed by the Fano factor [15],

$$F(T) = \sigma^2(T)/\mu(T). \quad (2)$$

We assume that for a weak stimuli, $\sigma(T)$ is not affected by the signal, and the Fano factor Eq. (2) and the discriminability Eq. (1) are related as

$$d' = \alpha\sqrt{\mu_0(T)/F_0(T)}, \quad (3)$$

where α is the relative change in the mean firing rate due to stimulus [7]. Equation (3) establishes the relation between spontaneous spike-count variability and discriminability of weak signals: smaller Fano factor, i.e., smaller variability of the spike count of spontaneous firing, results in larger discriminability. The dependence of the discriminability d' on the signal strength α is linear with the slope $\sqrt{\mu_0(T)/F_0(T)}$.

Variability of the spike count depends on the correlation structure of the spike train. The correlations can be quantified by the serial correlation coefficients $\rho_k = (\langle I_i I_{i+k}\rangle - \langle I\rangle^2)/\mathrm{var}(I)$ between interspike intervals, $I_i = t_i - t_{i-1}$. The asymptotic value of the Fano factor $F(T)$ is determined, if the serial correlation coefficients (SCCs) are summed over all lags: $\lim_{T\to\infty} F(T) = CV^2(1 + 2\Sigma_{k=1}^{\infty}\rho_k)$, where $CV = \sqrt{\mathrm{var}(I)}/\langle I\rangle$ is the coefficient of variation [16]. On intermediate time scales, the variability of the spike count depends crucially on the correlation structure, with positive ISI correlations acting to increase, while negative ISI correlations acting to suppress spike-count variability [7,17]. In the case of renewal spiking, when sequential ISIs are uncorrelated, $\rho_k = \delta_{0,k}$, the Fano factor approaches CV^2 for large T.

B. Model

A previous study [10] showed that long-lasting serial correlations are a generic result of a unidirectional coupling of two noisy oscillators. Here we aim to study how the signal detection performance of a neuron depends on the temporal structure of ISI correlations and to contrast the signal detection by a neuron with long-lasting serial ISI correlations versus a neuron lacking serial ISI correlations, when both neurons have identical first-order ISIs statistics. We employ two analytically tractable models, which we refer to as renewal and nonrenewal. These models differ in the coherence of driving noise, but otherwise have identical ISI distributions and linear-response properties. In both models, periodic spiking is mimicked by a perfect integrate-and-fire (PIF) model as in Refs. [12,13]. It is driven by two statistically independent Gaussian noises: a narrow-band noise $y(t)$, representing coherent stochastic oscillations, and an exponentially correlated noise $\eta(t)$, mimicking slow fluctuations [12]:

$$\dot{x} = \lambda + y + \eta,$$

$$\dot{y} = z, \quad \dot{z} = -\gamma z - \omega_0^2 y + A\omega_0\sqrt{2\gamma}\xi_1(t),$$

$$\dot{\eta} = -\eta/\tau_c + \sigma_\eta\sqrt{2/\tau_c}\xi_2(t). \quad (4)$$

The first equation for x in Eq. (4) is the perfect integrator. Whenever $x(t)$ reaches the threshold $x_b = 1$, a spike is generated and x immediately resets to the initial value $x_r = 0$. The parameter λ sets the mean firing rate of the model, $f_a = \lambda$. A constant stimulus is modeled as a change in the parameter λ. The two equations for y and z model narrow-band noise with the peak spectrum frequency $f_e = (1/2\pi)\sqrt{\omega_0^2 - \gamma^2/4}$ and the variance A^2. The coherence of the narrow-band noise is measured by the quality factor $Q = 2\pi f_e/\gamma$. Finally, the last equation describes the Ornstein-Uhlenbeck (OU) noise $\eta(t)$ with the correlation time τ_c and the variance σ_η^2. The OU noise is slow relative to the average period of spiking, $\tau_c \gg 1/\lambda$, and is introduced to account for slow fluctuations observed in sensory neurons [7]. The terms $\xi_{1,2}(t)$ are statistically independent Gaussian white noise sources with unit intensity.

For convenience, we introduce a vector $\mathbf{r}(t) = \{y(t), z(t), \eta(t)\}$ for the three-dimensional Gaussian process driving the perfect integrator. In the nonrenewal model, the noise vector $\mathbf{r}(t)$ evolves independently of spiking dynamics according to Eq. (4). In the renewal model, all components of the noise vector are reset at every spiking to a new value $\mathbf{r}_0 = \{y_0, z_0, \eta_0\}$ randomly sampled from the *noise upon firing* distribution $\mathcal{P}(\mathbf{r}_0)$. This is the stationary distribution of $\mathbf{r}$ values at the instants of firing in the nonrenewal model. In both models, the values of noise variables at the beginning of each ISI are drawn from the same distribution. Consequently, the ISI densities are identical for both models. However, the renewal model exhibits no ISI correlations, since the noise values at the beginning of each ISI are sampled from the distribution $\mathcal{P}(\mathbf{r}_0)$ at random, i.e., independently of the duration of the previous ISI and of the noise value at the instants of firing.

Figure 1 shows numerical simulations of the models. Although the firing patterns [Fig. 1(a)] of both renewal and nonrenewal models look very similar and the ISI densities are identical [Fig. 1(b)], their SCCs [Fig. 1(d)] are significantly different. This difference appears because the coherence of narrow-band noise $y(t)$ is destroyed in the renewal model as demonstrated in Fig. 1(c), where the power spectra of driving noise $y(t)$ in both models are depicted.

C. Analytical calculation of the Fano factor

Our goal is to contrast the discriminability in the renewal and nonrenewal models [Eq. (4)] and to analyze the discriminability in the nonrenewal model as a function of the coherence and frequency of the driving oscillations. To this end, we calculate analytically the Fano factor of the spike trains. Since the dependence of the discriminability d' on the signal strength α is linear [Eq. (3)], the relative performance of the renewal and nonrenewal models can be characterized by the ratio $R(T)$ of their discriminability slopes:

$$R(T) = \sqrt{\frac{\mu_n(T)}{F_n(T)}}\sqrt{\frac{F_r(T)}{\mu_r(T)}}. \quad (5)$$

The values of $R(T) > 1$ indicate that the signal detection performance is better in the nonrenewal than in the renewal model.

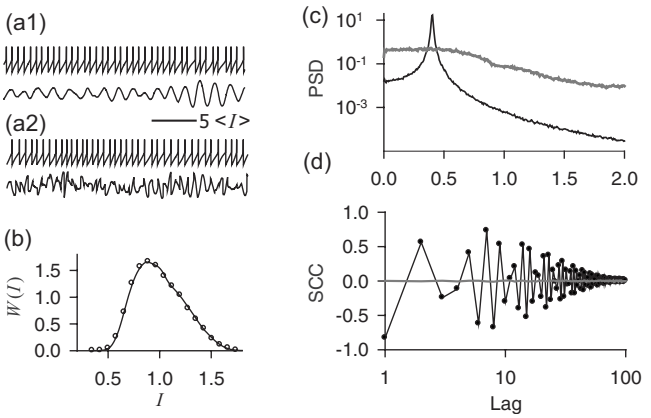


FIG. 1. Numerical simulation of model equations (4). (a): time series of the "voltage" variable x (upper trace) and noise $y(t)$ (lower trace) for the nonrenewal (a1) and renewal (a2) models. (b): probability density of ISIs, $W(I)$ for renewal (solid line) and nonrenewal (open circles) models. (c): power spectral densities (PSD) of noise $y(t)$ in renewal (gray line) and nonrenewal (black line) models. (d): SCCs of spike trains produced by renewal (gray line) and nonrenewal (black line with circles) models. The parameters are $\lambda=2$, $A^2=0.2$, $f_e/f_a=0.4$, $Q=20.0$, $\tau_c=375$, $\sigma_\eta=0.5\times10^{-4}$, $f_e/f_a=0.43$. Time axes are in units of mean interspike intervals. The frequency axis in panel (c) is scaled by the mean firing rate $f_a=2$.

Calculations of statistical properties of the spike train are largely simplified if N consecutive ISIs are replaced by a passage from x_r to sequentially $x_b, 2x_b, \ldots, kx_b \ldots$ and Nx_b [13]. Since the right-hand sides of Eq. (4) do not explicitly depend on x, this is possible if recrossing of a threshold is highly improbable, i.e., if the noise intensity is small relative to the drift term λ:

$$A^2 + \sigma_\eta^2 \ll \lambda^2/2. \tag{6}$$

Assume, we start our observation at time $t=0$, when the value of the voltage variable is x_0. The voltage value x_0+x at time T is determined by the transition probability density $P_x(x_0+x,T|x_0,0)=P_x(x,T|0,0)$ for the x variable:

$$P_x(x,T|0,0) = \int d\mathbf{r}d\mathbf{r}_0 P(x,\mathbf{r},T|0,\mathbf{r}_0,0)P_{st}(\mathbf{r}_0). \tag{7}$$

Here $P(x,\mathbf{r},T|x_0,\mathbf{r}_0,t_0)$ is the transition probability density for the multidimensional Gaussian process given by Eq. (4) without reset and $P_{st}(\mathbf{r}_0)$ is the stationary probability density of the noise variables. These probability densities are readily obtained analytically [18].

The number of spike counts N_T within the time window T is equal to the number of "thresholds" nx_b contained within the interval (x_0, x_0+x); see Fig. 2. For $nx_b<x<(n+1)x_b$, it holds

$$N_T = \begin{cases} n+1 & \text{with probability } \dfrac{x-nx_b}{x_b}, \\ n & \text{with probability } \dfrac{(n+1)x_b-x}{x_b}, \end{cases} \tag{8}$$

since the initial position x_0 is uniformly distributed along the x axis.

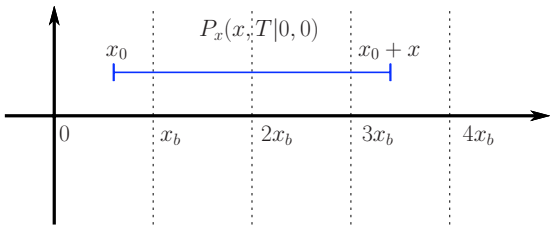


FIG. 2. (Color online) Illustration of the calculation of the Fano factor.

According to Eq. (8), the mean of the spike count $\langle N_{nx}\rangle$ for a fixed value of $x\in[nx_b,(n+1)x_b]$ is calculated as

$$\langle N_{nx}\rangle = (n+1)\frac{x-nx_b}{x_b} + n\frac{(n+1)x_b-x}{x_b} = x/x_b. \tag{9}$$

Analogously, the mean square of the spike count $\langle N_{nx}^2\rangle$ for a fixed value of $x\in[nx_b,(n+1)x_b]$ equals

$$\langle N_{nx}^2\rangle = (2n+1)\frac{x}{x_b} - n(n+1).$$

To obtain the mean $\mu_o(T)$ and the variance $\sigma_o^2(T)$ of the spike count, we average $\langle N_{nx}\rangle$ and $\langle N_{nx}^2\rangle$ over all values of x:

$$\mu_o(T) = \sum_{n=0}^{\infty}\int_{nx_b}^{(n+1)x_b}\langle N_{nx}\rangle P_x(x,T|0,0)dx,$$

$$\sigma_o^2(T) = \sum_{n=0}^{\infty}\int_{nx_b}^{(n+1)x_b}\langle N_{nx}^2\rangle P_x(x,T|0,0)dx - \mu_o^2(T). \tag{10}$$

This leads to the following expressions for the mean and the variance of the spike count:

$$\mu_o(T) = \frac{1}{x_b}\int_0^{\infty} xP_x(x,T|0,0)dx,$$

$$\sigma_o^2(T) = \sum_{k=0}^{\infty}\left[\frac{2k+1}{x_b}\int_{kx_b}^{(k+1)x_b} xP_x(x,T|0,0)dx - k(k+1)\int_{kx_b}^{(k+1)x_b} P_x(x,T|0,0)dx\right] - \mu_o^2(T). \tag{11}$$

The integrals in Eq. (11) are evaluated analytically and the infinite sum is approximated numerically.

For the renewal model, the mean of the spike count equals $\mu_r(T)=T\lambda$ and its variance is expressed as [15]

$$\sigma_r^2(T) = T\lambda + 2T\int_0^T\left(1-\frac{\tau}{T}\right)\rho^+(\tau)d\tau, \tag{12}$$

where $\rho^+(\tau)=\rho(\tau)-\lambda\delta(\tau)$ with $\rho(\tau)$ being the autocorrelation function of the renewal spike train [19]:

$$\rho^+(\tau) = \frac{\lambda}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}\frac{1-|\hat{W}(\omega)|^2}{|1-\hat{W}(\omega)|^2}d\omega - \lambda\delta(\tau). \tag{13}$$

Here, $\hat{W}(\omega)=\langle e^{i\omega I}\rangle$ is the Fourier transform of the ISI density $W(I)$. For weak noise, when recrossing of the threshold is

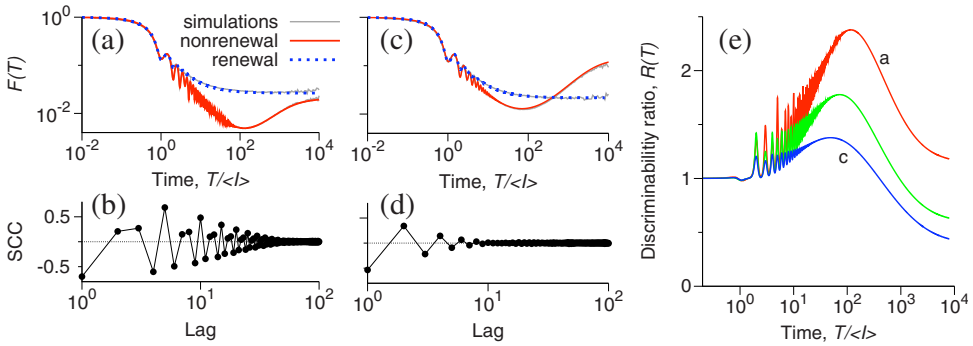


FIG. 3. (Color online) (a),(c): analytical Fano factor curves for renewal (dotted blue lines) and nonrenewal (solid red lines) models. For both models, the results of numerical simulations are shown as gray solid lines. (b),(d): numerically calculated SCCs ρ_k. The parameters are $\lambda=2$, $A^2=0.2$; (a),(b): $f_e/f_a=0.4$, $Q=20.0$, $\tau_c=375$, $\sigma_\eta=0.5\times10^{-4}$; (c),(d): $f_e/f_a=0.5$, $Q=4.0$, $\tau_c=900$, $\sigma_\eta=1.5\times10^{-4}$. (e): ratio of discriminability slopes in the nonrenewal and renewal models for parameter sets as in panels (a) and (c) (red and blue lines, respectively), and for $f_e/f_a=0.47$, $Q=10.0$, $\tau=600$, $\sigma_\eta=10^{-4}$ (green line).

very improbable, the ISI density can be approximated by the Rice density of level crossings [13,20,21]:

$$W(I)=\int d\mathbf{r}d\mathbf{r}_0(\lambda+y+\eta)P(x_b,\mathbf{r},I|x_r,\mathbf{r}_0,0)\mathcal{P}(\mathbf{r}_0). \quad (14)$$

Here $\mathcal{P}(\mathbf{r}_0)$ is related to the stationary density $P_{st}(\mathbf{r}_0)$ of the Gaussian noise via $\mathcal{P}(\mathbf{r}_0)=[(\lambda+y_0+\eta_0)/\lambda]P_{st}(\mathbf{r}_0)$ [16]. The integrals in Eq. (14) extend over the phase space region where $\dot{x}=\mu+\eta+y>0$. For weak noise intensities [see Eq. (6)], it holds that $(\lambda+y_0+\eta_0)/\lambda\approx1$ everywhere in the phase space, where $P_{st}(\mathbf{r}_0)$ is significantly different from zero. Therefore, without loss of accuracy, we approximate the noise upon firing density $\mathcal{P}(\mathbf{r}_0)$ in Eq. (14) by the stationary density $P_{st}(\mathbf{r}_0)$ and extend the integration limits to the whole phase space. The integrals in Eq. (14) are then evaluated analytically and then the Fano factor is calculated numerically using Eqs. (12) and (13) and the fast Fourier transform.

III. RESULTS AND DISCUSSION

The dependence of the Fano factor and SCCs on the coherence of oscillations is illustrated in Figs. 3(a)–3(d) for the nonrenewal (solid red lines) and renewal (dotted blue lines) models. The theoretical Fano curves perfectly overlap with the results of numerical simulations (gray lines). For intermediate time scales $10\langle I\rangle<T<200\langle I\rangle$, the Fano factor of the nonrenewal model, $F_n(T)$, becomes strikingly smaller than that of the renewal model, $F_r(T)$, and exhibits a minimum at $T\approx200\langle I\rangle$. $F_n(T)$ increases for $T>200\langle I\rangle$, due to slow fluctuations leading to positive ISI correlations [12,14], while $F_r(T)$ saturates to its theoretical limit of CV^2.

As expected, changes in Q almost do not influence the spike-count variability in the renewal model, but do have a strong impact on the variability in the nonrenewal model. With higher values of Q, corresponding to more coherent oscillations [Figs. 3(a) and 3(b), $Q=20$], serial correlations extend up to hundreds of ISIs, the minimum in the Fano factor of the nonrenewal model is deep, and the reduction in the spike-count variability relative to the renewal model can reach up to one order of magnitude, on time scales between 10 and $1000\langle I\rangle$. A decreased coherence of oscillations, reflected in lower Q values, results in brief ISI correlations, a shallower minimum in the Fano factor of the nonrenewal model, and a smaller reduction in variability relative to the renewal model [Figs. 3(c) and 3(d), $Q=4$]. The suppression of the spike-count variability due to oscillations translates to the enhancement of the discriminability, which is characterized by the discriminability ratio $R(T)$ [Eq. (5)] shown in Fig. 3(e). The maximal value of $R(T)$ decreases with decreasing oscillation coherence.

Figure 4 summarizes the dependence of discriminability on the oscillation to firing frequency ratio f_e/f_a and the quality factor Q. If f_e is low (small f_e/f_a), then several ISIs occur during one oscillation cycle, which results in positive ISI correlations on the time scale of the oscillation period [Fig. 4(d)]. For low Q values, these positive ISI correlations increase the variability in the nonrenewal model and lead to low discriminability and values of the slope ratio R less than 1. A Fano factor curve in this small-ratio low-Q regime is shown in Fig. 4(c) for $f_e/f_a=0.1$ and $Q=1$. For higher Q values, in addition to positive ISI correlations on shorter time scale, negative ISI correlations appear on the time scale of the coherence time of oscillations [Figs. 4(e) and 4(f), $f_e/f_a=0.1$, $Q=10$]. The variability in the nonrenewal model is reduced by these negative correlations and discriminability increases. If the oscillation is fast (large f_e/f_a), then several oscillation cycles occur during one ISI. The effect of the closed cycles on firing averages out through integration, hence all correlations are driven by the fraction of the remaining unclosed cycle. For fast oscillations, this is a brief time window that makes an ISI shorter or longer. Therefore the net modulation of each ISI is small and variability is almost unperturbed by oscillations: Fano factor curves in the nonrenewal and renewal models overlap largely, except for effects due to slow noise η [Fig. 4(g), $f_e/f_a=1.6$, $Q=80$]. However, since SCCs measure serial dependencies irrelative to the absolute ISI modulation, this almost negligible change in the variability is concurrent with strong extended negative ISI correlations [Fig. 4(h), cf. Figs. 3(a) and 3(b)]. Though, the discriminability is high in this large-ratio regime [Fig. 4(a)], a spike jitter would destroy all ISI correlations, increase CV and, hence, reduce the discriminability in both

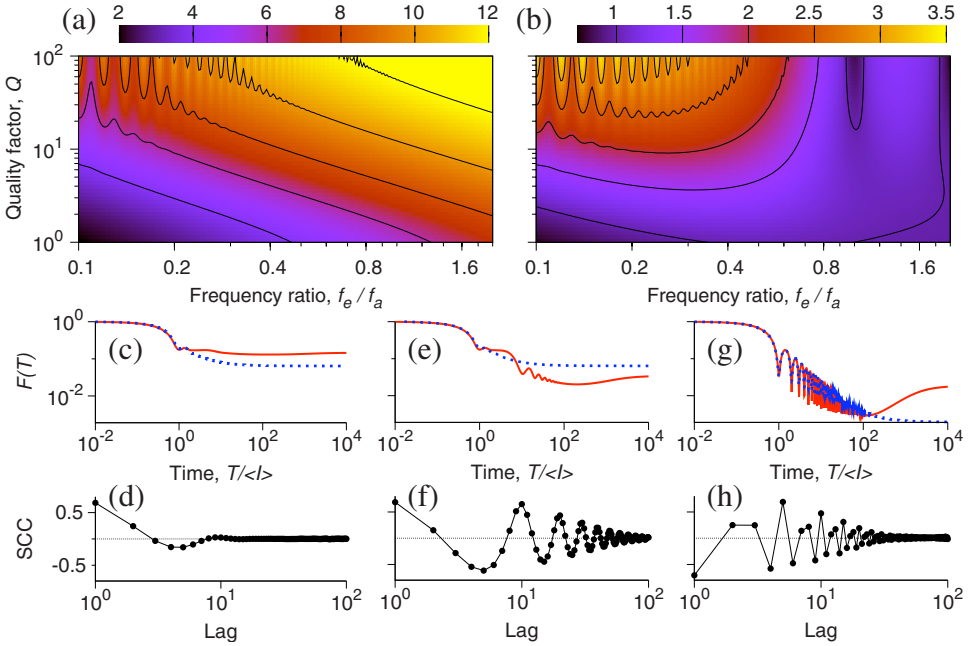


FIG. 4. (Color online) Discriminability in the nonrenewal model for 10% increase in the mean firing rate λ (a), and ratio of discriminability slopes in the nonrenewal and renewal models (b) at $T=50\langle I\rangle$: dependence on the frequency ratio f_e/f_a and on the quality factor Q of the driving oscillations. (c)–(h): SCCs and Fano factor curves for the renewal (dotted blue lines) and nonrenewal model (solid red lines). The parameters are (c),(d): $f_e/f_a=0.1$, $Q=1$; (e),(f): $f_e/f_a=0.1$, $Q=10$; (g),(h): $f_e/f_a=1.6$, $Q=80$. Other parameters are as in Fig. 3(a).

renewal and nonrenewal models. In contrast, ISI correlations in the near-0.5-ratio regime are robust to ISI jitters and the discriminability in the nonrenewal model is high even for large CV values.

We now show that analytical predictions of Fig. 4 can be used to explain the possible role of coherent epithelial oscillations in a particular sensory system, a paddlefish electroreceptor (PER). A PER can be represented as two unidirectionally coupled oscillators: the lumped epithelial oscillator driving the afferent oscillator [9,10]. Epithelial oscillation (EO) has a fundamental frequency $f_e\approx 26$ Hz, which is invariant under weak electric-field stimulations. Coherence of stochastic EO is reflected in high values of the experimentally measured quality factor Q in the range 10–20. Coherent EO can be represented by a narrow-band stochastic process [9]. The afferent neuron fires quasiperiodically at a mean rate f_a in the range 30–70 Hz [9]. Variability of afferent firing is mainly due to the input from EO [9]. At the same time, the EO induces long-lasting serial correlations of afferent ISIs [10].

The parameters of the nonrenewal model, Eq. (4), can be tuned to reproduce statistical characteristics of spontaneous dynamics of PERs including power spectrum, serial correlations, and Fano factor. As in analytical calculations, we contrast the signal detection by a PER driven by either coherent EO or by epithelial fluctuations with destroyed coherence, keeping ISIs distributions identical. Experimental abolition of EO is difficult to achieve without significant damage to the PER [9]. Instead, we mimicked the influence of a stimulus by insertion of spikes in a spontaneous spike train, recorded from a PER afferent, and in a surrogate renewal spike train obtained from the original by random shuffling of ISIs as in Ref. [4]. The first-order ISI statistics such as ISI density and coefficient of variation CV are identical for both spike trains. However, the shuffled spike train lacks any ISI correlations. The data were from a previous study on biperiodic oscillations in PERs [9]. Stationary segments of 120–1200 s of spontaneous activity recorded from 56 afferents from 19 paddlefish were used.

Figure 5(a) shows the Fano factor for a representative afferent firing spontaneously (solid line) and for the corresponding renewal surrogate (dotted line). As predicted by the analytic theory [cf. Fig. 3(a)], the Fano factor of the original spike train, $F_n(T)$, is significantly smaller than that of the renewal surrogate, $F_r(T)$, for $10\langle I\rangle<T<1000\langle I\rangle$, indicating suppressed variability in the original spike train.

We "stimulated" an afferent by inserting N spikes in random positions not already occupied by spontaneous spikes in nonoverlapping 1 s windows. This procedure increases the firing rate by N Hz, so that $\alpha=N\langle I\rangle$. Both the original spike train and the renewal surrogate yield a linear dependence of the discriminability d' on α described by Eq. (3) (data not shown). The slope of the discriminability at $T=1$ s was 2.1-fold steeper for the original than for renewal data: thus, there was twice better detection by the original spike train than by its renewal counterpart. The ratios of discriminability slopes $R(T)$ are shown in Fig. 5(c) for three different afferents and are in excellent agreement with analytical predictions [cf. Fig. 3(e)].

Figures 5(d) and 5(e) summarize results for the sample of 56 PER afferents. The mean firing rate in this population was 54.43 ± 9.54 Hz, with $CV=0.19\pm 0.05$. The epithelial-to-afferent frequency ratio $f_e/f_a=0.49\pm 0.07$. The value of the Fano factor was 0.031 ± 0.01 at $T=150$ ms and 0.010 ± 0.005 at $T=1$ s. The population average of the discriminability ratio R was 1.34 ± 0.13 at $T=150$ ms and 1.99 ± 0.35 at $T=1$ s indicating that discriminability was always better for a nonrenewal spike train. No significant dependence of R on CV or the frequency ratio was observed. These population results can be mapped to the parameter space of the model (Fig. 4). For the range of frequency ratios near 0.5, epithelial oscillations induce extended serial ISI anticorrelations and enhance the discriminability for realistic values of CV. This near-0.5-ratio high-Q regime corresponds

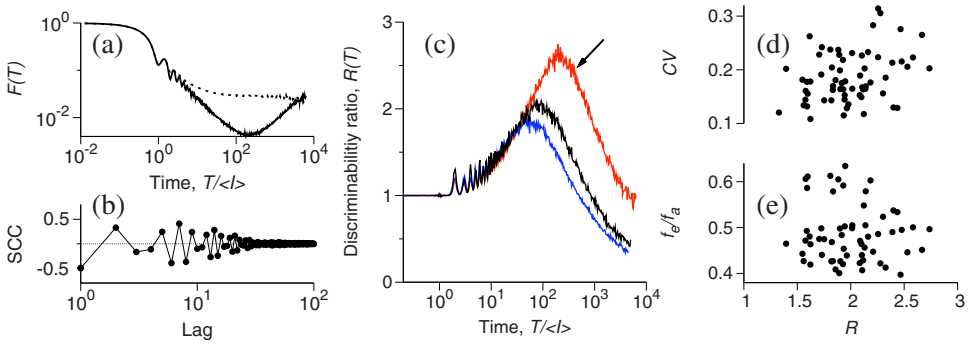


FIG. 5. (Color online) Experimental data from PER afferents. (a): Fano factors for the original (solid line) and renewal (dotted line) spike trains and (b): SCCs for an afferent with mean ISI $\langle I\rangle = 15.8$ ms, $CV = 0.163$, and $f_e/f_a = 0.43$. (c): the ratio of discriminability slopes, $R(T)$, for three different afferents. Arrow: afferent shown in panels (a) and (b). (d),(e): R at 1 s window vs CV and vs f_e/f_a for a population of 56 PER afferents.

to parameter values found in our PER database. The exact value of the frequency ratio is not crucial to achieve high discriminability in this regime. Hence PERs operate in a regime where reliable signal detection is possible without fine parameter tuning. Our analytical calculations agreed quantitatively with experimental data from PER revealing that coherent stochastic oscillations can enhance the discrimination performance of sensory neurons.

IV. CONCLUSION

We have studied the effect of stochastic oscillations acting as a source of external noise on discrimination of weak signals by periodically spiking neurons. Coherent stochastic oscillations lead to serial correlations extending to several dozens of interspike intervals, with the structure determined by the ratio of the frequency of driving stochastic oscillations (f_e) and the neuron mean firing rate (f_a) and by the quality factor (Q) of driving oscillations. This mechanism for generation of serial correlations is distinct from the one described for electroreceptors of weakly electric fish [4], where negative ISI correlations appear presumably due to a negative feedback from adaptation ionic currents or modulation of firing threshold [7,8] and do not extend beyond few interspike intervals. We demonstrate that spike-count variability induced by these oscillations can be suppressed relative to a corresponding renewal model, if the frequency of driving oscillations is smaller than the mean firing rate and their coherence is large enough. Suppressed spike-count variability translates to an enhanced discrimination capacity of the neuron characterized in this study by the discriminability measure d'. Our analytical calculations yield parameter range where discriminability is enhanced by coherent stochastic oscillations.

We applied these analytical predictions to study the role of epithelial oscillations in peripheral electroreceptors of paddlefish. The population averages of the ratio of epithelial to sensory neuron frequencies and of the quality factor of epithelial oscillations are nicely mapped on the analytically derived parameter range of discriminability enhancement. This gives a strong argument in favor that epithelial oscillations serve to enhance discriminability of weak signals.

In conclusion, our analytical calculations underscore that coherent oscillations can improve discrimination capacity of spiking neurons in a feed-forward scenario. Such mechanism may be relevant to a broad range of phenomena involving unidirectionally coupled oscillators, e.g., to spiking neurons receiving feed-forward oscillatory input from a network of CNS neurons or possibly to low-frequency sensors of electric and magnetic fields [22,23].

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Grant No. DC04922 and by the Biomimetic Nanoscience and Nanotechnology program of Ohio University and by Deutsche Forschungsgemeinschaft Grant No. SFB 555.

[1] G. Buzsaki, *Rhythms of the Brain* (Oxford University Press, New York, 2006).
[2] A. T. Schaefer, K. Angelo, H. Spors, and T. W. Margrie, PLoS Biol. **4**, e163 (2006).
[3] N. Masuda and B. Doiron, PLOS Comput. Biol. **3**, e236 (2007).
[4] R. Ratnam and M. E. Nelson, J. Neurosci. **20**, 6672 (2000).
[5] M. J. Chacron, L. Maler, and J. Bastian, Nat. Neurosci. **8**, 673 (2005).
[6] D. J. Mar, C. C. Chow, W. Gerstner, R. W. Adams, and J. J. Collins, Proc. Natl. Acad. Sci. U.S.A. **96**, 10450 (1999).
[7] M. J. Chacron, A. Longtin, and L. Maler, J. Neurosci. **21**, 5328 (2001).
[8] M. J. Chacron, B. Lindner, and A. Longtin, Phys. Rev. Lett. **92**, 080601 (2004).
[9] A. B. Neiman and D. F. Russell, J. Neurophysiol. **92**, 492 (2004).
[10] A. B. Neiman and D. F. Russell, Phys. Rev. E **71**, 061915

(2005).
[11] I. Fuwape and A. B. Neiman, Phys. Rev. E **78**, 051922 (2008).
[12] J. W. Middleton, M. J. Chacron, B. Lindner, and A. Longtin, Phys. Rev. E **68**, 021920 (2003).
[13] B. Lindner, Phys. Rev. E **69**, 022901 (2004).
[14] T. Schwalger and L. Schimansky-Geier, Phys. Rev. E **77**, 031914 (2008).
[15] F. Gabbiani and C. Koch, in *Computational Neuroscience*, edited by C. Koch and I. Segev (MIT Press, Cambridge, MA, 1998).
[16] D. R. Cox, *The Statistical Analysis of Series of Events* (Wiley, New York, 1966).
[17] S. B. Lowen and M. C. Teich, J. Acoust. Soc. Am. **92**, 803 (1992).
[18] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1996).
[19] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1963).
[20] S. O. Rice, Bell Syst. Tech. J. **24**, 51 (1945).
[21] T. Verechtchaguina, I. M. Sokolov, and L. Schimansky-Geier, Phys. Rev. E **73**, 031108 (2006).
[22] J. A. Acebrón, W.-J. Rappel, and A. R. Bulsara, Fluct. Noise Lett. **3**, L341 (2003).
[23] A. R. Bulsara, V. In, A. Kho, A. Palacios, P. Longhini, J. Neff, G. Anderson, C. Obra, S. Baglio, and B. Ando, Meas. Sci. Technol. **19**, 075203 (2008).