Behavioral/Cognitive

Collective Activity of Many Bistable Assemblies Reproduces Characteristic Dynamics of Multistable Perception

Robin Cao,[1,3] **Alexander Pastukhov,**[4] **Maurizio Mattia,**[3] and **Jochen Braun**[1,2]

[1]Institute of Biology, Otto-von-Guericke University, 39120 Magdeburg, Germany, [2]Center for Behavioral Brain Sciences, 39120 Magdeburg, Germany, [3]Istituto Superiore di Sanità, 00161 Rome, Italy, and [4]Institute of Psychology, Otto-Friedrich-University, 96047 Bamberg, Germany

The timing of perceptual decisions depends on both deterministic and stochastic factors, as the gradual accumulation of sensory evidence (deterministic) is contaminated by sensory and/or internal noise (stochastic). When human observers view multistable visual displays, successive episodes of stochastic accumulation culminate in repeated reversals of visual appearance. Treating reversal timing as a "first-passage time" problem, we ask how the observed timing densities constrain the underlying stochastic accumulation. Importantly, mean reversal times (i.e., deterministic factors) differ enormously between displays/observers/stimulation levels, whereas the variance and skewness of reversal times (i.e., stochastic factors) keep characteristic proportions of the mean. What sort of stochastic process could reproduce this highly consistent "scaling property?" Here we show that the collective activity of a finite population of bistable units (i.e., a generalized Ehrenfest process) quantitatively reproduces all aspects of the scaling property of multistable phenomena, in contrast to other processes under consideration (Poisson, Wiener, or Ornstein-Uhlenbeck process). The postulated units express the spontaneous dynamics of attractor assemblies transitioning between distinct activity states. Plausible candidates are cortical columns, or clusters of columns, as they are preferentially connected and spontaneously explore a restricted repertoire of activity states. Our findings suggests that perceptual representations are granular, probabilistic, and operate far from equilibrium, thereby offering a suitable substrate for statistical inference.

Key words: attractor cell assemblies; birth-death process; cortical columns; first-passage time; multistable perception; scaling property

---

Significance Statement

Spontaneous reversals of high-level perception, so-called multistable perception, conform to highly consistent and characteristic statistics, constraining plausible neural representations. We show that the observed perceptual dynamics would be reproduced quantitatively by a finite population of distinct neural assemblies, each with locally bistable activity, operating far from the collective equilibrium (generalized Ehrenfest process). Such a representation would be consistent with the intrinsic stochastic dynamics of neocortical activity, which is dominated by preferentially connected assemblies, such as cortical columns or clusters of columns. We predict that local neuron assemblies will express bistable dynamics, with spontaneous active-inactive transitions, whenever they contribute to high-level perception.

Introduction

Response time distributions reveal much about the neural mechanisms underlying perceptual decisions (Ratcliff and Smith,

2004; Smith and Ratcliff, 2009). A gradual and noisy accumulation of sensory information, which triggers a choice response when exceeding a criterion amount, explains the respective response time distributions of correct and error responses as well as their dependence on task factors, such as choice utility or the desired speed or accuracy of response. In several instances, neurophysiological studies were able to confirm and extend this quantitative understanding (Schall, 2001; Sugrue et al., 2005; Gold and Shadlen, 2007), revealing distinct cortical substrates for accumulating and for evaluating sensory information (Resulaj et al., 2009; Shadlen and Kiani, 2013).

Here we extend this approach to the multistable appearance of ambiguous visual displays (Leopold and Logothetis, 1999; Blake and Logothetis, 2002; Sterzer et al., 2009). Multistable sit-

uations are characterized by spontaneous perceptual decisions in the form of occasional, sudden reversals of visual appearance (Leopold and Logothetis, 1999; Braun and Mattia, 2010). The timing of such reversals depends primarily on stimulus factors (Levelt, 1967; Brascamp et al., 2006) and only secondarily on cognitive or task factors (e.g., Pearson et al., 2008).

Even though a dominant appearance is perceived as stable until a reversal transpires, its robustness gradually declines, as revealed by perturbation experiments (Wolfe, 1984; Nawrot and Blake, 1989; Petersik, 2002; Kang and Blake, 2010). It is thought that support for the currently dominant appearance degrades while, simultaneously, support for an alternative appearance accumulates, until a critical differential is reached (Laing and Chow, 2002; Wilson, 2003; Moreno-Bote et al., 2007; Noest et al., 2007; Hohwy et al., 2008). The cortical sites associated with, respectively, accumulating and evaluating competing evidence from multistable displays (e.g., Knapen et al., 2011) appear to overlap with the sites identified for choice tasks (Resulaj et al., 2009; Shadlen and Kiani, 2013).

The reversal times of ambiguous displays have been characterized in exquisite detail. Astonishingly, the distribution of reversal times maintains a characteristic, gamma-like shape (Levelt, 1967; Blake et al., 1971; Walker, 1975; De Marco et al., 1977; Murata et al., 2003; Pastukhov and Braun, 2007), even though mean reversal times vary widely between observers, displays, and input levels (Fox and Herrmann, 1967; Borsellino et al., 1972; Walker, 1975; Zhou et al., 2004; Brascamp et al., 2005). Moreover, the distribution in question is considerably more variable ("wider") and less skewed ("shorter-tailed") than typical choice time distributions (Carpenter, 2012; Pearson et al., 2014). The constancy of distribution shape constitutes a "scaling property", as it implies that higher moments scale as appropriate powers of the mean.

Here we ask the following: what sort of stochastic process could reproduce both the scaling property and the short-tailed skewness of reversal time distributions? What could guarantee, for all displays and observers, that gradual accumulation of sensory information leading to perceptual reversals is contaminated by proportional noise: slower accumulations by less, and faster accumulations by more noise? And what sort of process could ensure that reversal time distributions are consistently short-tailed?

As a possible solution to this conundrum, we propose a finite population of cortical columns or clusters that are individually bistable in that they transition spontaneously between active and inactive states (generalized birth-death or generalized Ehrenfest [GE] process) (Karlin and McGregor, 1965; van Kampen, 1981; Cao et al., 2014). The discreteness of such a representation would automatically keep normalized moments constant and ensure a scaling property. Moreover, the asymmetric fluctuations of a finite population would ensure a short-tailed distribution. The proposed representation mimics the intrinsic stochastic dynamics of neocortical activity, which is dominated by preferentially connected local assemblies, such as cortical columns or clusters of columns.

Materials and Methods

Perceptual observations
Observers. Procedures were in accordance with the Declaration of Helsinki and were approved by the medical ethics board of the Otto-von-Guericke University (Magdeburg, Germany). All participants had normal or corrected-to-normal vision, were naive to the purpose of experiments, and were paid for participating.

Apparatus and general procedure. Stimuli were generated with MATLAB (The MathWorks). Observers responded using a keyboard. Background luminance was kept at 36 cd/m^2. The experimental room

was lit dimly (ambient luminance at 80 cd/m^2). For visual multistability, we considered situations in which strong retinal inputs are spontaneously suppressed: motion-induced-blindness (MIB), binocular rivalry (BR) (Campbell and Howell, 1972; Leopold and Logothetis, 1999; Bonneh et al., 2001), spontaneous reversals of illusory rotation in depth (kinetic depth effect [KDE]) (Wallach and O'Connell, 1953; Sperling and Dosher, 1994), spontaneous shifts in the apparent direction of motion (moving plaids [MP]) (Adelson and Movshon, 1982; von Grünau and Dubé, 1993), and spontaneous changes in the 3D appearance of line drawings (Necker cube [NC]) (Meng and Tong, 2004). The datasets on KDE, NC, and MP, as well as one of six datasets on BR, were published previously (Pastukhov and Braun, 2007; Pastukhov et al., 2013). New datasets were collected on MIB and on BR.

For auditory multistability, we considered spontaneous transitions between auditory segregation and integration (auditory streaming) (Bregman, 1994; Winkler et al., 2012). This dataset was kindly provided by I. Winkler and S. Denham (personal communication).

For choice reaction tasks, we reanalyzed published information on saccadic reaction times (saccade reaction time) (Carpenter, 2012) and on working memory retrieval times (memory reaction time) (Pearson et al., 2014). Both studies report timing densities for individual observers.

Binocular rivalry. Six observers participated in the experiment (4 male, 2 female). Stimuli were displayed on an LCD screen (EIZO ColorEdge CG303W, resolution 2560 × 1600 pixels, viewing distance 104 cm, single pixel subtended 0.014°, refresh rate 60 Hz) and were viewed through a mirror stereoscope. Chin and head rests were used to stabilize viewing position.

Two grayscale circular orthogonally oriented gratings (45° and −45°) were presented foveally to each eye. Gratings had diameter of 1.6° (spatial period 2 cyc/deg). To avoid a sharp edge, grating contrast was modulated with Gaussian envelope (starting inner radius 0.6°, $\sigma = 0.2°$). Tilt and phase of gratings were randomized for each block. Five contrast levels were used as follows: 6.25%, 12.5%, 25%, 50%, and 100%. Contrast of each grating was systematically manipulated, so that each contrast pair was presented in two blocks (50 blocks in total). Each block was 2 min long and separated by a compulsory 1 min break. Observers reported on the tilt of the visible grating by continuously pressing one of two arrow keys. They were instructed to press only during exclusive visibility of one of the gratings, so that mixed percepts were indicated by neither key being pressed (25 ± 8% of total presentation time). To facilitate binocular fusion, gratings were surrounded by a dichoptically presented square frame (outer size 9.8°, inner size 2.8°).

Motion-induced blindness. Twenty observers participated in the experiment (12 male, 8 female). Stimuli were presented on a CRT screen (Iiyama VisionMaster Pro 514, www.iiyama.com; resolution 1600 × 1200 pixels, refresh rate 100 Hz). The viewing distance was 73 cm so that each pixel subtended ∼0.019°. Target was a yellow circle (diameter 0.2°), presented 1° above the fixation. Mask rotated at 1 Hz and consisted of 8 × 8 grid of crosses (arm length 0.6°, intercross distance 0.15°). Observers reported episodes of target disappearance by keeping the space key pressed for the entire duration of the episode. Blocks lasted 1 min and were separated by a compulsory 30 s. In total, 32 experimental blocks were measured for each observer.

Scaling property
Definitions and relevant quantities. A distribution of samples t is characterized by moments of the distribution (density). Distribution shape (Table 1) may be quantified in terms of the mean $\mu_1 \equiv \langle t \rangle$ and the central moments $\mu_2 \equiv \langle (t - \mu_1)^2 \rangle$, $\mu_3 \equiv \langle (t - \mu_1)^3 \rangle$, etc., or, equivalently, in terms of normalized moments, such as the coefficient of variation $c_v = \mu_2^{1/2}/\mu_1$ and the skewness $\gamma_1 = \mu_3/\mu_2^{3/2}$.

A scaling property obtains if central moments are proportional to corresponding powers of the mean or, equivalently, if normalized moments are constant as follows:

$$\mu_2 \propto \mu_1^2, \quad \mu_3 \propto \mu_1^3 \Leftrightarrow c_v = const, \quad \gamma_1 = const$$

Table 1. Comparison of investigated random-walk processes[a]

Name	Type	v^*_{drift}	v^*_{noise}	Scaling property	Distribution shape $c_v \approx 0.6$ & $\gamma_1 \approx 2c_v$
Wiener	Gaussian diffusion	$\dfrac{x_{in}}{\tau}$	$\dfrac{\sigma^2}{\tau}$	—	—
Modified Wiener	Gaussian diffusion	$\dfrac{x_{in}}{\tau}$	$\dfrac{x_{in}}{\tau\alpha}$	√	—
Balanced Poisson	Infinite birth-death	$v_E - v_I$	$v_E + v_I$	√	—
Ornstein-Uhlenbeck	Gaussian diffusion	$\dfrac{x_{in}}{\tau}$	$\dfrac{\sigma^2}{\tau}$	—	—
Modified OU	Gaussian diffusion	$\dfrac{x_{in}}{\tau}$	$\dfrac{x_{in}}{\tau\alpha}$	√	—
Generalized Ehrenfest	Finite birth-death	v_+	$\dfrac{v_+}{N}$	√	√
Continuous Generalized Ehrenfest	Gaussian diffusion	v_+	$\dfrac{v_+}{N}$	√	—

[a]Normalized moments remain constant with input and a scaling property is obtained (in the drift-dominated regime), when accumulation rate v^*_{drift} and dispersion rate v^*_{noise} increase with input in the same proportions. The experimentally observed skewness is uniquely expressed by a generalized Ehrenfest process.

Denoting individual realizations of activity with $x(t)$ and the ensemble average with $\langle x(t) \rangle$, the accumulation and dispersion rates are defined as the respective time-derivatives of the mean and variance as follows:

$$v_{drift} \equiv \frac{d}{dt}\langle x(t) \rangle, \quad v_{noise} \equiv \frac{d}{dt}[\langle x^2(t) \rangle - \langle x(t) \rangle^2]$$

In the general case, these rates will be both input- and state-dependent. In other words, they will change with both sensory input s and instantaneous activity $x(t)$ as follows:

$$v_{drift} = v_{drift}(x, s), \quad v_{noise} = v_{noise}(x, s)$$

If v_{drift} increases with input and activity $x(t)$ traverses a range $x_0 \leq x(t) \leq \theta$, the mean first-passage time (FPT) $\mu_1 = \langle t_{fp} \rangle$ is expected to decrease with input s and increase with threshold θ (compare Fig. 2), consistent with experimental observations. The same is true for the central moments μ_2 and μ_3.

When activity $x(t)$ remains far below an input-dependent steady-state $x_{in}(s)$ (drift-dominated regime), the accumulation and dispersion rates are independent of $x(t)$ and therefore stationary as follows:

$$v_{drift}(s, x) \xrightarrow{x \ll x_{in}} v^*_{drift}(s), \quad v_{noise}(s, x) \xrightarrow{x \ll x_{in}} v^*_{noise}(s)$$

Together with threshold θ, the stationary rates v^*_{drift} and v^*_{noise} are the minimal ingredients for capturing the observed statistics of reversal times.

Sufficient condition for a scaling property. We now formulate a sufficient condition between θ, v^*_{drift}, and v^*_{noise} for coefficients of variation of FPTs to remain constant with input. This condition holds for any drift-dominated accumulation ($x \ll x_{in}$; see Fig. 2A) in which the variance of FPTs $\sigma_t^2 = \langle t^2 \rangle - \langle t \rangle^2$ is proportional to variance of activity $\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2$ at time $\langle t \rangle$, so that the respective coefficients of variation are approximately the same (Wiener, Ornstein-Uhlenbeck, Poisson, or Ehrenfest processes) as follows:

$$\sigma_t \propto \sigma_x \quad \Rightarrow \quad c_v^{(t)} \approx c_v^{(x)}$$

Experimental evidence shows that increased input decreases the mean time $\langle t \rangle$ of first passing a threshold θ decreases. It follows that increased input must increase accumulation rates v^*_{drift} as follows:

$$\langle t \rangle = \theta / v^*_{drift}, \quad v^*_{drift} = f(x_{in}), \quad \partial f / \partial x_{in} > 0$$

If v^*_{noise} remains constant with input, it further follows that the variance at $\langle t \rangle$ decreases with v^*_{drift} and the coefficient of variation at $\langle t \rangle$ decreases with the square root of v^*_{drift} as follows:

$$\sigma_x^2 = v^*_{noise} \langle t \rangle = v^*_{noise} \theta / v^*_{drift} \propto f^{-1}(x_{in}),$$

$$c_v^{(x)} = \frac{\sigma_x}{\theta} = \sqrt{v^*_{noise} / \theta v^*_{drift}} \propto f^{-1/2}(x_{in})$$

so that no scaling property is obtained.

However, if both v^*_{noise} and v^*_{drift} increase with input in the same proportions, the coefficient of variation remains constant as follows:

$$v^*_{noise} \propto v^*_{drift} \propto f(x_{in}) \quad \Rightarrow \quad c_v \propto \sqrt{v^*_{noise} / \theta v^*_{drift}} = const$$

and the necessary condition for a scaling property is satisfied.

Stochastic processes under study

We summarize the salient characteristics of the stochastic processes under discussion, in particular, the moments of the FPT distributions. Where explicit expressions are unavailable, we resorted to a numerical integration of the related Langevin equation with standard first-order methods (Risken, 1984). For each combination of x_{in} and θ, we simulated 10^5 FPTs in time-steps of 0.01 s.

Balanced Poisson (BP) process

This process approximates excitatory and inhibitory inputs incrementing and decrementing the membrane potential of a neuron (Tuckwell, 1988). Excitatory (respectively inhibitory) spikes arrive with a fixed rate v_E (respectively v_I), so that either spike counts form independent Poisson processes $N_E(t)$ (respectively $N_I(t)$). FPT is defined as time t from initial count $N_E(0) = N_I(0) = 0$ to a threshold count θ. The distribution of FPT has mean $\mu_1 = \langle t \rangle$, coefficient of variation c_v, and skewness γ_1 (Tuckwell, 1988) as follows:

$$\mu_1 = \frac{\theta}{v_E - v_I}, \quad c_v = \sqrt{\frac{v_E + v_I}{\theta(v_E - v_I)}}, \quad \gamma_1 = 3\sqrt{\frac{v_E + v_I}{\theta(v_E - v_I)}} = 3c_v$$

Discrete Ehrenfest process

This generalized birth-death process is used in many contexts, including statistical mechanics, chemical physics, and theoretical biology, to model the statistical dynamics of macroscopic ensembles composed of discrete microscopic units (Karlin and McGregor, 1965; van Kampen, 1981; Emch and Liu, 2002; Ghosh et al., 2006; Cao et al., 2014). We consider a population of N stochastic and bistable units, each transitioning spontaneously and independently between inactive and active states. Transition rates v_+ (activation) and v_- (inactivation) are assumed to be stationary. If n is the number of units active at a given time (with $0 \leq n \leq N$), the overall rate of upward transitions within the population is $(N - n)v_+$, and the rate of downward transitions is $n v_-$. Because of this state dependence, the step distribution of the discrete Ehrenfest process is asymmetric, especially when n approaches the extremes of its range: $n \approx 0$ or $n \approx N$. Moments of the FPT density may be obtained from the renewal equation as finite sums over Krawtchouk polynomials, as detailed previously (Cao et al., 2014).

Wiener process
A Wiener process combines constant (but input-dependent) accumulation with constant dispersion and produces a noisy linear accumulation (Cox and Miller, 1972). It satisfies a Langevin equation with constant infinitesimal drift and variance,

$$\tau\dot{x} = x_{in} + \sqrt{\tau}\sigma\xi(t)$$

and normally distributed "white noise" $\xi(t)$ is defined as follows:

$$\langle\xi(t)\rangle = 0, \qquad \langle\xi(t)\,\xi(t')\rangle = \delta(t - t')$$

From initial condition $x_0 = 0$, the distribution of FPT to a threshold θ is an inverse Gaussian or Wald distribution (Tuckwell, 1988), with moments as follows:

$$\mu_1 = \frac{\theta\tau}{x_{in}}, \qquad c_v = \frac{\sigma}{\sqrt{x_{in}\,\theta}}, \qquad \gamma_1 = \frac{3\sigma}{\sqrt{x_{in}\,\theta}} = 3c_v$$

Ornstein-Uhlenbeck process
An Ornstein-Uhlenbeck process combines a state-dependent (as well as input-dependent) accumulation with constant dispersion, resulting in a noisy exponential relaxation to a steady-state state x_{ss} with zero drift (Cox and Miller, 1972; Risken, 1984). It satisfies a Langevin equation with state- and input-dependent infinitesimal drift, and constant infinitesimal variance as follows:

$$\tau\dot{x} = -x + x_{in} + \sqrt{\tau}\sigma\xi(t), \qquad x_{ss} = x_{in}$$

Moments of the FPT distribution have been derived in terms of infinite series (Inoue et al., 1995), or in terms of nested integrals (Brunel, 2000).

Continuous Ehrenfest process
In the continuous limit, discrete random walks may be approximated by Gaussian diffusion processes (Cox and Miller, 1972; Risken, 1984). For a GE process, the continuous limit is a Cox-Ingersoll-Ross process (Cox et al., 1985) in which both infinitesimal drift and variance are state- and input-dependent (van Kampen, 1981) as follows:

$$\tau\dot{x} = -x + x_{in} + \sqrt{\tau}\sqrt{\frac{x_{in} - \beta x(t)}{N}}\xi(t)$$

where $\beta = \frac{\nu_+ - \nu_-}{\nu_+ + \nu_-}$ is proportional to the relative rate difference. The step distribution of this process is Gaussian and therefore symmetric.

Escape process
To model instantaneous escape across an adapting threshold, we assume that the instantaneous escape probability reflects normally distributed noise with mean 0 and variance σ^2. Specifically, we compute the instantaneous probability that the noise exceeds the distance to the (time-varying) threshold $\theta(t)$ as follows:

$$p_{esc}(t) = \frac{\nu}{\sqrt{2\pi}} \int_{(x_{ss} - \theta(t))/\sigma}^{\infty} e^{-x'/2}\,dx'$$

where ν is an escape rate. Dividing time into discrete intervals Δt, we approximate the FPT density $f(t)$ as follows:

$$\ln f(t) = \ln\left[p_{esc}(n\Delta t)\right] + \sum_{i=0}^{n-1} \ln\left[1 - p_{esc}(i\Delta t)\,\Delta t\right],$$

and evaluated the moments numerically.

Overview
Spiking neuron simulations. To illustrate a possible neural realization of a GE process, we simulated "strongly coupled" assemblies of excitatory and inhibitory neurons (see Fig. 8A,B), with each assembly expressing a spontaneous and independent bistable dynamics (Mattia et al., 2013). Each such assembly comprised 125 excitatory leaky-integrate-and-fire neurons ("foreground"), which were weakly coupled to an additional 875 excitatory neurons ("background"), as well as 250 inhibitory neu-

rons. The background population is not strictly necessary and included here only for the sake of verisimilitude. The connection probability between any two neurons was $c = 80\%$. Excitatory synaptic efficacy between "foreground" neurons, "background" neurons, and between the two was $J_{fore} = 0.618$ mV, $J_{back} = 0.438$ mV, and $J_{inter} = 0.402$ mV, respectively. Inhibitory synaptic efficacy was $J_I = -1.50$ mV, and the efficacy of excitatory synapses onto inhibitory neurons was $J_{IE} = 0.560$ mV. Finally, "foreground" neurons, "background" neurons, and "inhibitory" neurons each received independent Poisson spike trains of 2340, 2280, and 2280 Hz, respectively. Other settings were as in Mattia et al. (2013). As a result of these settings, "foreground" activity transitioned spontaneously between a low state of ~3 Hz and a high state of ~40 Hz.

Weakly coupled assemblies (see Fig. 8C) were obtained by reducing synaptic efficacies $J_{fore} = 0.566$ mV and $J_{back} = 0.431$ mV, as well as increasing efficacies $J_{inter} = 0.409$ mV (to maintain overall level of activity). For "foreground" neurons (each with external Poisson inputs of 2400 Hz), the firing rate was ~3 Hz. To reproduce the gradual accumulation of activity by strongly coupled assemblies, we increased (at a suitable pace) external Poisson inputs to 2760 Hz, eventually raising the firing rate to ~30 Hz.

To analyze the spiking activity of "foreground" neurons, we established the instantaneous average rate $r(t)$ and generated a equally sized population of "surrogate" neurons as nonhomogeneous Poisson point processes with the same rate $r(t)$. The statistics of "foreground" neuron spiking could then be compared with that of "surrogate" neurons (see Fig. 8D,E).

Formal treatment of stochastic processes
Discrete and continuous random walks require different formal treatments, namely, a Master Equation or a Fokker-Planck-equation, respectively. We provide a brief synopsis of these approaches, to justify, or at least indicate the origin of, the various results cited in the main text. A full treatment may be found in Tuckwell (1988) and Risken (1984).

Discrete random walks
Continuous-time random walks in a discrete state-space K constitute a subclass of Markov processes and their individual realizations form lattice paths. Such processes are governed by a set of gain-loss relations on the probability of each state $n \in \mathcal{K}$, called the Master Equation as follows:

$$\frac{dP_n(t)}{dt} = \sum_{k\in\mathcal{K},k\neq n} \{W_{k\to n}P_k(t) - W_{n\to k}P_n(t)\} \qquad (1)$$

where W denotes the transition rate from one state to another.

BP process. Given transition rates ν_E, ν_I, the Master Equation reads as follows:

$$\frac{dP_n(t)}{dt} = \nu_E P_{n-1}(t) - \nu_I P_{n+1}(t) \qquad n \in \mathbb{Z}.$$

This can be solved analytically by means of the generating function $G(z, t) = \sum_k P_k(t)\,z^k$, assuming initial conditions $P_n(t = 0) = \delta_{0n}$ as follows:

$$P_n(t) = \left(\frac{\nu_E}{\nu_I}\right)^{\frac{n}{2}} e^{-(\nu_E + \nu_I)t}\,I_n(2t\sqrt{\nu_E\nu_I})$$

illustrated in Figure 7F, where

$$I_n(x) \equiv \sum_{k=0}^{+\infty} \frac{1}{k!(k + n)!}\left(\frac{x}{2}\right)^{2k+n}$$

is the modified Bessel function of the first kind.

GE process. Given transition rates ν_+, ν_-, the Master Equation reads as follows:

$$\frac{dP_n(t)}{dt} = [N - (n - 1)]\nu_+ \cdot P_{n-1}(t) + (n + 1)\nu_- \cdot P_{n+1}(t)$$
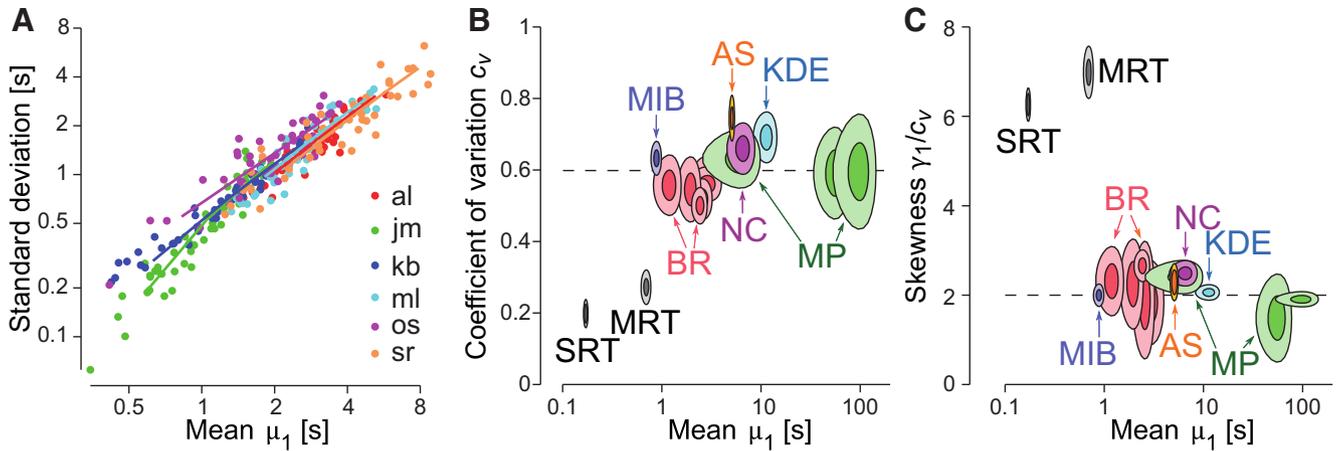
$$- [(N - n)\nu_+ + n\nu_-] \cdot P_n(t)$$

Figure 1. Scaling property of multistable perception. **A**, SD and mean μ_1 of the distribution of dominance durations. Individual values (round symbols) for 25 contrast combinations of BR and 6 observers (different colors). Proportionality is maintained both across conditions (contrast combinations) and across observers. Linear regression (colored lines) yields a highly consistent $c_v = 0.66 \pm 0.04$ (mean $\pm$ SD). **B**, **C**, Normalized moments for the distribution of dominance durations observed with different multistable perceptions: auditory streaming (AS), binocular rivalry (BR), kinetic-depth effect (KDE), motion-induced blindness (MIB), moving plaids (MP), and Necker cube (NC). Several stimulus conditions are shown for BR and MP. For comparison, saccade reaction time (SRT) and memory reaction time (MRT) are included as well. **B**, Coefficient of variation c_v, as a function of the mean μ_1. All multistable situations exhibit comparable variability, with $c_v \approx 0.6$ (dashed line). **C**, Skewness γ_1, in multiples of c_v, as a function of the mean μ_1. All multistable situations exhibit comparable skewness, with $\gamma \approx 2c_v$ (dashed line). **B**, **C**, Shaded areas represent 1 and 2 SEM across observers. The size of the SEM ellipses reflects the number of observers in each case, which ranged from 3 to 59 (see Materials and Methods).

For initial condition, $P_n(t = 0) = \delta_{0n}$, the Master Equation may be solved as before to obtain the following:

$$P_n(t) = \binom{N}{n} [p(1 - e^{-t/\tau_{in}})]^n$$

$$\times [pe^{-t/\tau_{in}} + q]^{N-n} \xrightarrow{t \to \infty} \binom{N}{n} p^n q^{N-n}$$

where $\tau_{in} = \dfrac{1}{\nu_+ + \nu_-}$, $p = \nu_+ \tau_{in} = x_{in}$, and $q = 1 - p = \nu_- \tau_{in}$. The time evolution of this distribution is illustrated in Figure 7E.

Continuous random walks
Gaussian random walks in a continuous state-space can derive from the continuous limits of discrete state-space diffusion processes (Cox and Miller, 1972). Typically, continuous processes are more tractable, both analytically and numerically, than their discrete counterparts. The time-dependent probability density $p(x, t)$ of a continuous random variable $x(t)$ is governed by a Fokker-Planck equation (Risken, 1984) as follows:

$$\tau \frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}(\mu(x, x_{in}) \, p(x,t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(\sigma^2(x, x_{in}) \, p(x,t))$$

where $\mu(x, x_{in})$ and $\sigma(x, x_{in})$ represent "accumulation" or "drift" and "dispersion" or "noise," respectively. In general, both terms may depend on the current state x and on external input x_{in}. For instance, in the continuous limit of the GE process, accumulation and dispersion rates reflect the input-dependent individual transition rates as well as the number of units available for activation or inactivation. The total rates of increments ρ_+ and decrements ρ_- (see Fig. 7A, B) are as follows:

$$\rho_+ = N \nu_+ (1 - x) = N x_{in} (1 - x)/\tau_{in},$$

$$\rho_- = N \nu_- x = N (1 - x_{in}) x / \tau_{in},$$

where x is instantaneous activity (fraction of active units), and are shown in Fig. 7C, D. Therefore, the infinitesimal drift and variance are as follows:

$$\mu(x, x_{in}) = \frac{1}{N}(\rho_+ - \rho_-) = \frac{x_{in} - x}{\tau_{in}}$$

$$\sigma^2(x, x_{in}) = \frac{1}{N^2}(\rho_+ + \rho_-) = \frac{x_{in} - \beta x}{N}, \qquad \beta = \frac{\nu_+ - \nu_-}{\nu_+ + \nu_-}$$

and are shown in Figure 7C, D.

The dynamics of single realizations of such a process are described by a stochastic differential equation (Langevin equation) as follows:

$$\tau \dot{x} = \mu(x, x_{in}) + \sqrt{\tau} \, \sigma(x, x_{in}) \, \xi(t)$$

where $\xi(t)$ is normally distributed "white noise." The Langevin formulation makes clear that the drift term describes a deterministic force that is identical on every trial, whereas the noise term captures stochastic forces that differ from trial to trial.

Results

To assess how well (or poorly) multistable phenomena conform to a scaling property, we obtained and compared observations from six canonical multistable situations (five visual, one auditory). We also considered two choice reaction tasks as examples of "normal" vision.

Wide range of mean dominance periods
In multistable phenomena, mean dominance periods vary widely with stimulus strength, observer, and display type. A particularly well-studied situation is BR, where both left and right eye contrast may be varied (Levelt, 1967; Moreno-Bote et al., 2010). A representative set of observations is summarized in Figure 1A. For a given observer, mean dominance periods varied fourfold between 25 contrast combinations (a 5 × 5 matrix of left and right eye contrasts; see Materials and Methods). Importantly, mean dominance periods decreased with stimulus contrast. Between observers, mean dominance periods differed by a further factor of 4. Between different display types, mean dominance periods ranged even more widely, as illustrated in Figure 1B, C. More than six types of multistable visual displays (BR, KDE, MIB, MB, NC) and one multistable auditory scene (auditory streaming), mean dominance periods varied over 2 orders of magnitude, from <1 s to >100 s (see Materials and Methods).

Compared with mean dominance periods, choice reaction times are far less variable. A twofold range with stimulus strengths or between observers appears typical (Carpenter, 2012; Pearson et al.,

2014). The two choice reaction tasks illustrated here span the full range of human reaction times, from 100 *ms* to 1 s (Fig. 1*B*,*C*).

Invariant shape of dominance distribution

The intuitive definition of a scaling property is that distributions exhibit identical shape when scaled to their respective distribution mean. Formally, a scaling property obtains if central moments μ_2 and μ_3 are proportional to corresponding powers of the mean μ_1 as follows:

$$\mu_2 \propto \mu_1^2, \qquad \mu_3 \propto \mu_1^3$$

or, equivalently, if normalized moments, such as the coefficient of variation c_v or the skewness γ_1 are constant (see Materials and Methods) as follows:

$$c_v = const, \qquad \gamma_1 = const$$

Multistable phenomena maintain a remarkably strict scaling property over different stimulus strengths, observers, and display types. For example, the SD of the dominance distribution is (nearly) proportional to the mean, as illustrated by the linear regression lines in Figure 1*A*.

To assess the generality of this finding, we established the mean, SD, coefficient of variation, and skewness of the distribution of dominance periods for different visual and auditory conditions: auditory streaming, MIB, NC, and KDE, BR for stimulus contrasts $c \in \{0.06, 0.12, 0.25, 0.5, 0.6, 1.0\}$, and MP for continuous display, intermittent display, and intermittent display without attention (Pastukhov and Braun, 2007). Figure 1*B*, *C* illustrates the results in terms of the mean and SEM across observers. The coefficient of variation remained consistently near $c_v \approx 0.6$ (Fig. 1*B*) and the skewness remained consistently near $\gamma_1 \approx 2c_v$ (Fig. 1*C*), the value corresponding to a Gamma distribution. In other words, a strict scaling property was maintained over all investigated situations.

Choice reaction times do not conform this pattern. Distribution means are less variable and distribution shapes combine smaller variance (($c_v \approx 0.2$) with greater skewness ($\gamma > 6c_v$).

To summarize, the dominance distribution of multistable phenomena exhibits a scaling property with three aspects: (1) distribution shape (normalized moments) is highly preserved across conditions (stimulus strengths, observers, and display types); (2) distribution variance is comparatively large ($c_v \approx 0.6$) and distribution skewness comparatively small ($\gamma_1 \approx 2c_v$), corresponding to a gamma distribution; and (3) distribution mean varies widely across conditions and decreases with stimulus strength.

Dominance durations as FPTs

Our aim is to compare the ability of various stochastic processes to reproduce the scaling property of multistable perception. An FPT framework (Fig. 2) provides convenient and concise terms for this purpose. In this framework, the timing of perceptual events (such as a choice response or a perceptual reversal) is interpreted as the first passage of a threshold by some underlying neural activity. Specifically, this activity is assumed to grow from some initial level toward the threshold level due to both deterministic and stochastic factors (i.e., factors that remain the same and that change randomly, respectively, from trial to trial).

FPT models account well for response time densities in many task situations (Ratcliff and Smith, 2004; Smith and Ratcliff, 2009; Carpenter, 2012; Pearson et al., 2014). They should not be taken literally in the sense that activity necessarily accumulates in a single neural population. More likely, sensory information
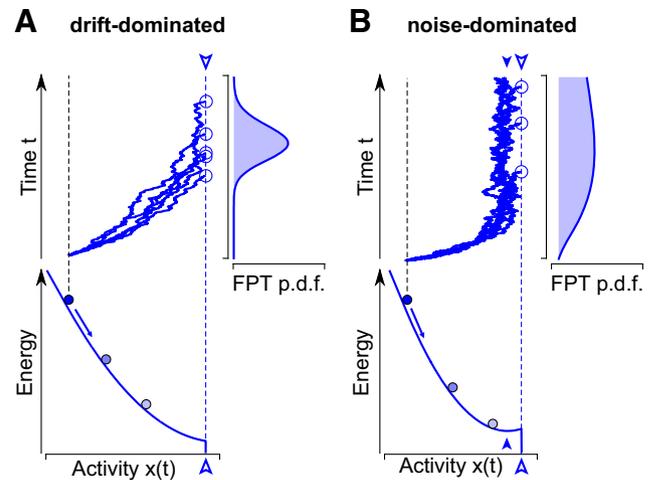

Figure 2. FPTs of a threshold level θ by stochastic neuronal activity $x(t)$. In this framework, a perceptual decision or reversal is triggered when $x(t) \geq \theta$ for the first time. *A*, In a drift-dominated regimen, deterministic forces drive activity to θ (open arrow) and beyond. *B*, In a noise-dominated regimen, deterministic forces drive activity merely to a steady-state (filled arrow), some distance below θ (open arrow). Thus, θ may be reached only with the help of stochastic forces (noise). *A*, *B*, In both regimens, individual realizations of neural activity $x(t)$ develop from an initial level (black dashed line) to a threshold level (blue dashed line, open arrows). Because of stochastic factors, every realization reaches threshold at a different time (open circles). This variability results in a probability distribution of FPT, which is illustrated on the right. Deterministic forces may be visualized in terms of an energy landscape, which is shown below. The sign of the energy gradient at threshold (open arrow) distinguishes drift- and noise-dominated regimens.

accumulates concurrently in several competing populations, so that a threshold is reached when the differential activity exceeds a certain value (e.g., Usher and McClelland, 2001; Smith and Ratcliff, 2009; Bundesen et al., 2015).

In the case of multistable reversals, the available evidence supports a differential threshold in that the currently dominant representation appears to degrade while, concurrently, the currently suppressed representation appears to recover (Wolfe, 1984; Nawrot and Blake, 1989; Petersik, 2002; Kang and Blake, 2010). One conceivable scenario is that gradated representations at sensory levels interact with categorical representations at perceptual levels. Perhaps differential activity at sensory levels progressively diverges from differential activity at perceptual levels, until the latter becomes destabilized and dominance reverses. This temporarily realigns the two levels and the cycle begins anew. However, here we are not concerned with such details. We simply assume that dominance periods reflect the time at which stochastically accumulating (differential) activity exceeds a fixed threshold for the first time, in other words, that reversal times represent FPTs of one sort or another.

FPT models may operate in two distinct regimens, illustrated in Figure 2. In drift-dominated regimens, deterministic forces are sufficient to drive activity to threshold. Stochastic forces merely introduce some variability (Fig. 2*A*). In noise-dominated regimens, deterministic forces drive activity toward a steady-state, which lies some distance *below* threshold (Fig. 2*B*). Stochastic forces are needed to drive activity *away* from steady-state (filled arrow), such as to cover the remaining distance to threshold (open arrow). The difference is best appreciated in an effective energy landscape visualization: the energy gradient takes opposite sign at threshold (Fig. 2*A*,*B*, open arrows).

It is intuitive that FPT density depends sensitively on both deterministic and stochastic forces over the entire activity

range traversed. A change to any single variable, deterministic force, stochastic force, or traversed range, will alter the FPT density and all of its moments. In general, first and higher moments will not change in a proportional manner. The following sections will show that two or more variables must change in a proportional manner for a scaling property to come about.

Comparison of stochastic processes
Stochastic processes may be compared in terms of the rates with which they deterministically "accumulate" and stochastically "disperse" activity over time (for definition, see Materials and Methods). In general, accumulation rate $\nu_{drift}(s, x)$ and dispersion rate $\nu_{noise}(s, x)$ will depend both on sensory input s and on instantaneous activity $x(t)$. Together with initial activity and threshold activity θ, these quantities determine the statistics of reversal times.

The following sections compare various stochastic processes in search of a parsimonious explanation for a scaling property. We begin with current models of perceptual decisions and multistable reversals, which use continuous stochastic processes with constant noise (Wiener process, Ornstein-Uhlenbeck process). Next, we modify these models with input-dependent noise. Finally, we consider discrete stochastic processes (Poisson process, Ehrenfest process), where deterministic and stochastic forces share a common physical origin.

Continuous stochastic processes with constant noise
Perceptual decisions are commonly modeled in terms of diffusion-like processes (Smith, 2000; Ratcliff and Smith, 2004; Smith and Ratcliff, 2009; Smith et al., 2014). Whereas mean accumulation (drift) is assumed to vary with input, stochastic dispersion (noise) and threshold level are assumed to be constant (Wiener process). The possibility that accumulation varies with activity and vanishes for a particular steady-state level (Ornstein Uhlenbeck process) has also been considered (Smith, 2000; Smith et al., 2014). Multistable reversals are widely modeled in similar terms but additionally invoke an adapting threshold (Laing and Chow, 2002; Moreno-Bote et al., 2007; Noest et al., 2007; Wilson, 2007; Shpiro et al., 2009; Moreno-Bote et al., 2010; Seely and Chow, 2011). We now consider each of these possibilities in turn.

In a Wiener process, the accumulation rate changes with sensory input s, but the dispersion rate (noise) remains constant as follows:

$$\nu_{drift}^* = \frac{x_{in}(s)}{\tau}, \qquad \nu_{noise}^* = \frac{\sigma^2}{\tau}, \qquad x_0 \le x(t) \le \theta \quad (2)$$

where τ is a characteristic time and σ is noise amplitude. Drift must be positive, and the development of $x(t)$ is always drift-dominated (Fig. 2A). The moments of the FTP depend on $x_{in}(s)$ and threshold θ and are illustrated in Figure 3A (middle columns). Whereas the mean μ_1 shortens with x_{in} and lengthens with threshold θ, the coefficient of variation c_v decreases with both input x_{in} and θ. The skewness is consistently high with $\gamma_1 = 3c_v$. The experimentally observed scaling property is not reproduced (Fig. 3A, rightmost column).

In an Ornstein-Uhlenbeck process, the infinitesimal drift varies with activity and vanishes for $x(t) = x_{in}$ as follows:

$$\nu_{drift}^* = \frac{x_{in}(s)}{\tau}, \qquad \nu_{noise}^* = \frac{\sigma^2}{\tau}, \qquad x_0 \le x(t) \le \theta \quad (3)$$

The threshold-crossing of $x(t)$ is drift-dominated when $x_{in} > \theta$, but noise-dominated when $x_{in} < \theta$ (Fig. 2). In both regimens, FPT moments may be obtained analytically (Inoue et al., 1995; Brunel, 2000) or numerically (Fig. 3B, middle columns). A scaling property is reproduced neither by noise-dominated nor by drift-dominated regimens. In both cases, the coefficient of variation c_v varies with x_{in} and skewness $\gamma_1 \approx 3c_v$ is too high.

Finally, current models of multistable perception additionally assume a time-dependent threshold $\theta(t)$ that progressively decreases with a characteristic time τ_A. Assuming that $x(t)$ fluctuates spontaneously about a constant mean and that a reversal occurs instantaneously whenever $x(t) > \theta(t)$, the moments of FPT may be obtained analytically (Molini et al., 2011) or by numerical simulation (Fig. 3C, middle columns). Similar to the Ornstein-Uhlenbeck process, both noise-dominated and drift-dominated regimens are obtained. In both regimens, the coefficient of variation c_v varies with input.

Interestingly, the skewness ranges from $\gamma_1 \approx 3c_v$ in the noise-dominated regimen to $\gamma_1 \approx c_v$ in the drift-dominated regimen. The experimentally observed skewness of $\gamma_1 \approx 2c_v$ obtains only for one particular combination of of x_{in} and θ (Fig. 3C, middle columns, arrow). This confirms and extends previous reports that a precise balance between adaptation and noise is required to attain the characteristic, gamma-like distribution of multistable reversals (van Ee, 2009; Pastukhov et al., 2013).

To summarize, if the accumulation rate ν_{drift}^* varies with input while the dispersion rate ν_{noise}^* remains constant, the normalized FPT moments vary with input and a scaling property is not obtained. This conclusion holds independently of details, such as the presence or absence of a steady-state or of an adapting threshold.

Continuous stochastic processes with input-dependent noise
The previous section showed that a scaling property cannot be obtained if only accumulation rate, but not dispersion rate, changes with input. Accordingly, we now consider scenarios where accumulation and dispersion change proportionally with sensory input s.

A modified Wiener process with input-dependent dispersion (noise) is as follows:

$$\nu_{drift}^* = \frac{x_{in}(s)}{\tau}, \qquad \nu_{noise}^* = \frac{x_{in}(s)}{\alpha\tau}, \qquad x_0 \le x(t) \le \theta$$

$$(4)$$

has the FPT moments illustrated in Figure 4A (middle columns). As expected, the normalized moments c_v and γ_1 remain constant with x_{in}, whereas mean FPT spans a wide range of values. The constant α determines the requisite threshold precision. The higher its value, the narrower the range of thresholds over which a given range of c_v is obtained. However, the skewness is $\gamma_1 \approx 3c_v$ higher than experimentally observed (Fig. 4A, rightmost column).

A modified Ornstein-Uhlenbeck process with input-dependent dispersion (noise) (Cox et al., 1985) is as follows:

$$\nu_{drift}^* = \frac{x_{in}(s)}{\tau}, \qquad \nu_{noise}^* = \frac{x_{in}(s)}{\alpha\tau}, \qquad x_0 \le x(t) \le \theta$$

$$(5)$$

exhibits the FPT moments illustrated in Figure 4B (middle columns). In the drift-dominated regimen, the FPT density is modulated similarly as in a Wiener process. In the noise-dominated regimen, the FPT density resembles that of the origi-
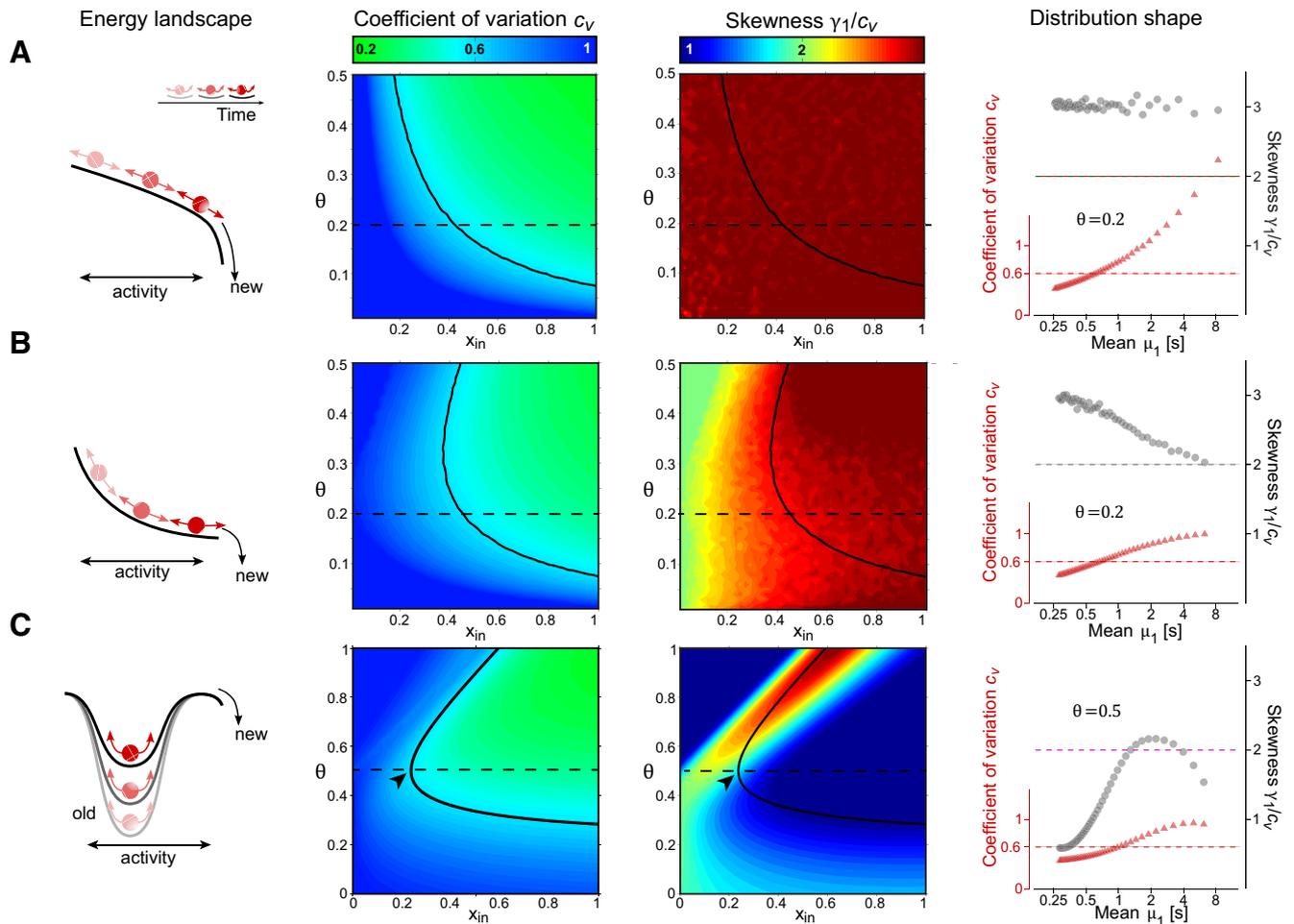
Figure 3. Continuous stochastic processes with constant noise: normalized moments of FPT density as functions of input x_{in} and threshold θ (both in units of σ). Leftmost column, Energy landscape and evolution over time (schematic). Middle left and right columns, Coefficient of variation c_v and skewness γ_1 (in units of c_v). Black curves represent $c_v = 0.6$. Black arrows indicate the distribution shape of multistable perception ($c_v = 0.6$ and $\gamma_1 = 2c_v$). Dashed line indicates a particular choice of threshold θ. Rightmost column, FPT distribution shape as function of mean μ_1 (format as in Fig. 1B, C). For constant θ (value given by inset), μ_1 decreases as x_{in} increases from 0 to 1. **A**, Input-dependent drift and constant noise (Wiener process). **B**, Input- and activity-dependent drift and constant noise (Ornstein-Uhlenbeck process). **C**, Adaptation of threshold and instantaneous escape (current models of multistable perception).

nal, unmodified Ornstein-Uhlenbeck process. As before, skewness $\gamma_1 \approx 3c_v$ is consistently higher than experimentally observed (Fig. 4B, rightmost column).

To summarize, maintaining the proportionality $\nu^*_{noise} \propto \nu^*_{drift}$ ensures the constancy of normalized moment in drift-dominated regimes. However, the experimentally observed skewness of $\gamma_1 \approx 2c_v$ cannot be reproduced. The reason is that all processes with symmetric noise, so-called Gaussian diffusion processes, are characterized by an inverse Gaussian or Wald distribution of FPT with $\gamma_1 \approx 3c_v$ (Tuckwell, 1988).

Processes with discrete stochastic events

The two previous sections showed that a scaling property obtains when sensory input affects accumulation and dispersion proportionally. A natural way to ensure this proportionality are processes comprising discrete and spontaneous events, such as Poisson events, simply because event rate will affect both accumulation and dispersion. Several levels of neural activity may be described in terms of Poisson events (e.g., spikes, up-state and down-state transitions, avalanches). We now consider two complementary examples: a BP process and a GE process.

A BP process combines two independent Poisson processes contributing, respectively, activity increments and decrements of unit size with homogeneous rates ν_E and ν_I Tuckwell (1988). Several realizations $n(t)$ of such a process are illustrated in Figure 5A. In the illustrated example, the initial distribution is deterministic with $p(n, t = 0) = \delta(n - 4)$ and $\nu_E = 4.8$ Hz is comfortably larger than $\nu_I = 3.2$ Hz, ensuring that mean and variance of $x(t)$ increase linearly over time. The evolution of $p(n, t)$ over time is illustrated in Figure 7E.

In general, the accumulation and dispersion rates are as follows:

$$\nu^*_{drift} = \nu_E - \nu_I, \qquad \nu^*_{noise} = \nu_E + \nu_I, \qquad n \leq \theta_n \quad (6)$$

and a scaling property is guaranteed (Fig. 6A), provided sensory input s changes the event rates in a proportional manner, so that the proportionality between accumulation and dispersion is maintained as follows:

$$\nu_E(s) \propto \nu_I(s) \propto f(s) \qquad \Rightarrow \qquad \nu^*_{drift} \propto \nu^*_{noise}$$

However, the skewness $\gamma_1 = 3c_v$ is once again larger than experimentally observed. The reason is that a BP process remains a

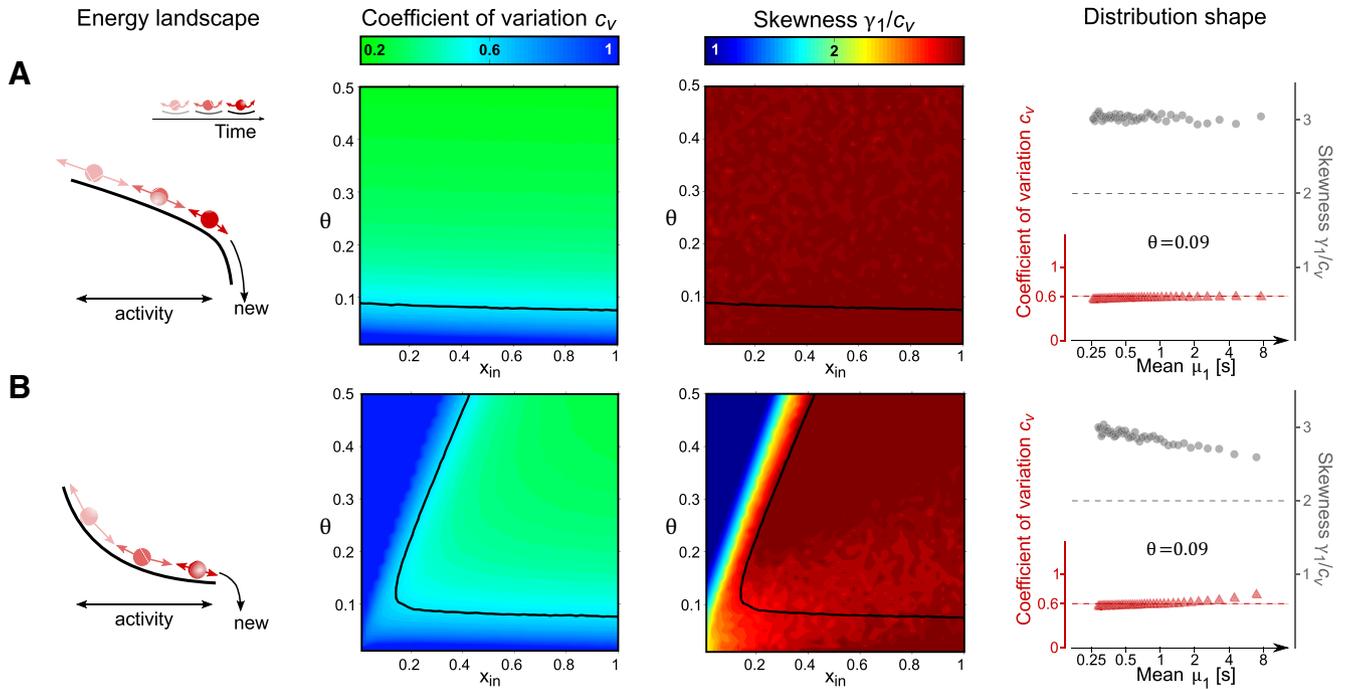**Figure 4.** Continuous stochastic processes with input-dependent noise: normalized moments of FPT density as functions of input x_{in} and threshold θ. Leftmost column, Energy landscape and evolution over time (schematic). Middle left and right columns, Coefficient of variation c_v and skewness γ_1 (in units of c_v). Black curves represent $c_v = 0.6$. Rightmost column, FPT distribution shape as function of mean μ_1 (format as in Fig. 1B, C). For constant θ (value given by inset), μ_1 decreases as x_{in} increases from 0 to 1. **A**, Input-dependent drift and noise (modified Wiener process). **B**, Input- and activity-dependent drift and noise (modified Ornstein-Uhlenbeck process).

diffusive process (despite its discrete nature), in that activity disperses symmetrically, as shown in Figure 7F.

We now compare and contrast a GE process, in which discrete increments and decrements of activity originate within a finite population of spontaneously bistable units (Fig. 5B–D). Such a population may become saturated or depleted, so that activation (inactivation) rates will vary with the number of inactive (active) units. This renders both accumulation and dispersion rates dependent on current activation x, with important consequences for FPT moments. Originally introduced in 1907 as the famous "dog-flea" model of diffusion, this generalized birth-death process provides a minimal model for the statistical dynamics of many microscopic and mesoscopic systems (Karlin and McGregor, 1965; van Kampen, 1981; Emch and Liu, 2002; Ghosh et al., 2006; Cao et al., 2014).

We consider a population of N bistable units, each of which transitions spontaneously between inactive and active states. The transitions are independent Poisson events with rates $\nu_+ = \nu_+(s)$ (activation) and $\nu_- = \nu_-(s)$ (inactivation), set by sensory input s. Additionally, we assume that the ratio ν_+/ν_- grows monotonically with sensory input s as follows:

$$1 - x \underset{\nu_-}{\overset{\nu_+}{\rightleftharpoons}} x, \qquad d/ds(\nu_+/\nu_-) > 0.$$

On average, $x(t)$ relaxes exponentially to a steady-state activation x_{in} with a characteristic time τ_{in}, both of which change with sensory input s as follows:

$$x_{in}(s) = \frac{\nu_+}{\nu_+ + \nu_-}, \qquad \tau_{in}(s) = \frac{1}{\nu_+ + \nu_-}.$$

This simple behavior of the average is deceptive, however, for the detailed dynamics is considerably more complex. Several realizations $x(t)$ are illustrated in Figure 5B. Starting from a deterministic initial state ($n_0 = 4$), mean and variance of $x(t)$ increase linearly over time, as $\nu_+ = 0.06$ Hz is comfortably larger than $\nu_- = 0.04$ Hz. The realizations of the GE process remain more compact than those of the BP in Figure 5A, even though the total increment and decrement rates are identical ($\nu_+ N = 4.8$ Hz and $\nu_- N = 3.2$ Hz). The difference is even more evident in the evolution of activity $p(x, t)$ over time. Comparing the GE process (Fig. 7F) with the BP process (Fig. 7E), two qualitative differences are apparent: the GE accumulation is faster and the GE dispersion is asymmetric. We now consider the reasons for these important differences.

In a GE process, accumulation and dispersion rates reflect the input-dependent transition rates as well as the number of units available for activation or inactivation. As the two populations are typically of unequal size, increment and decrement rates may differ considerably. The resulting infinitesimal drift and variance (see Materials and Methods) are as follows:

$$\nu^*_{drift} = \nu_+(s), \qquad \nu^*_{noise} = \frac{\nu_+(s)}{N}$$

and are shown in Figure 7C, D. At low activation levels $x \ll 0.5$, increments are far more frequent than decrements, explaining the two qualitative features, fast accumulation and asymmetric dispersion, mentioned above.

Next, we turn to the moments of the FPT density and their dependence on x_{in} and θ. Given an activity range $x_0 \leq x(t) \leq \theta$, FPT moments may be obtained analytically (Cao et al., 2014) and are illustrated in Figure 6B (middle columns). In the drift-dominated regime $x_{in} > \theta$ (indicated by arrows), normalized moments c_v and γ_1 remain nearly constant with changing input x_{in}. Moreover, the skewness asymptotes to the experimentally observed value $\gamma_1 \approx 2c_v$ (Fig. 6B, rightmost column). Only a small distance $\theta - x_0$ between threshold and initial condition is

required to obtain $c_v \approx 0.6$. These conclusions hold for any initial condition and threshold in the low-activity regimen $x \ll 0.5$ where the coefficient of variation is approximately (Cao et al., 2014) the following:

$$c_v \simeq 1/\sqrt{(\theta - x_0)\, N}$$

In the limit of large x_{in}, the FPT density of a GE process approaches an exact gamma distribution, with skewness $\gamma_1 = 2c_v$ and excess kurtosis $\gamma_2 = 6\, c_v^2$.

To summarize, a GE process operating far from steady-state ($x \ll x_{in}$) reproduces the scalar property of multistable perception for two reasons. First, both accumulation and drift increase proportionally with x_{in}, ensuring nearly constant normalized moments c_v and γ_1 (Fig. 7C,D). Second, positive accumulation is strong and negative accumulation negligible (Fig. 7A,B). This entails an asymmetric (non-Gaussian) dispersion that "thins out" realizations lingering near x_0 and keeps the activity distribution compact and asymmetric (Fig. 7E). The comparatively compact dispersion of activity then translates into a short-tailed FPT distribution with skewness $\gamma_1 \approx 2c_v$.

Finally, to conclusively demonstrate the importance of asymmetric noise, we compare the discrete GE process with its continuous equivalent, a Cox-Ingersoll-Ross process where the infinitesimal drift and variance $\mu(x, x_{in})$ and $\sigma^2(x, x_{in})$, are both input- and state-dependent. In the drift-dominated regimen, accumulation and dispersion rates are as follows:

$$\nu^*_{drift} = \nu_+(s), \qquad \nu^*_{noise} = \frac{\nu_+(s)}{N},$$

$$0 \leq x(t) \leq \theta \quad (7)$$

Even though this process matches both the accumulation rates ν^*_{drift} and the dispersion rates ν^*_{noise} of the discrete GE process, it is a Gaussian diffusive process and operates with symmetric noise. The moments of the FPT density are illustrated in Fig. 6C (middle columns). As with the original, discrete process, both c_v and γ_1 remain constant with changing x_{in}. However, unlike the discrete process, the skewness asymptotes to the larger value $\gamma_1 \approx 3c_v$ of an inverse Gaussian distribution (Fig. 6C, rightmost column).

We conclude by listing the necessary and sufficient conditions for reproducing the scaling property of multistable perception as follows: (1) a population of N discrete, locally bistable units that spontaneously activate and inactivate with rates ν_+ and ν_-, respectively; (2) the relative rate ν_+/ν_-, and thus also the steady-state activation $x_{in} = \nu_+/(\nu_- + \nu_+)$, grows with sensory input s; (3) both initial and threshold activation are comparatively small and far from steady-state, $x \ll x_{in}$; (4) population size N is limited by threshold precision; and (5) to obtain the experimentally observed coefficients of variation $c_v \approx 0.6$, the distance between threshold and initial condition must be small, $\theta - x_0 \approx 0.05$.

Neuronal realization of GE process

A neuronal realization of collective activity accumulating in a population of locally bistable assemblies is illustrated in Figure 8. We simulated 20 assemblies of strongly coupled spiking neurons, each balancing excitation and inhibition, such as to express a bistable attractor dynamics. Activations and inactivations occurred spontaneously, driven by endogeneous activity fluctuations (finite-size noise), and both transition rates changed with external input (for details, see Materials and Methods). To initiate a stochastic accumulation of collective activity, we abruptly altered external input, such as to raise the steady-state fraction of active assemblies from near zero to near unity. The spike raster in Figure 8A shows individual assemblies activating at
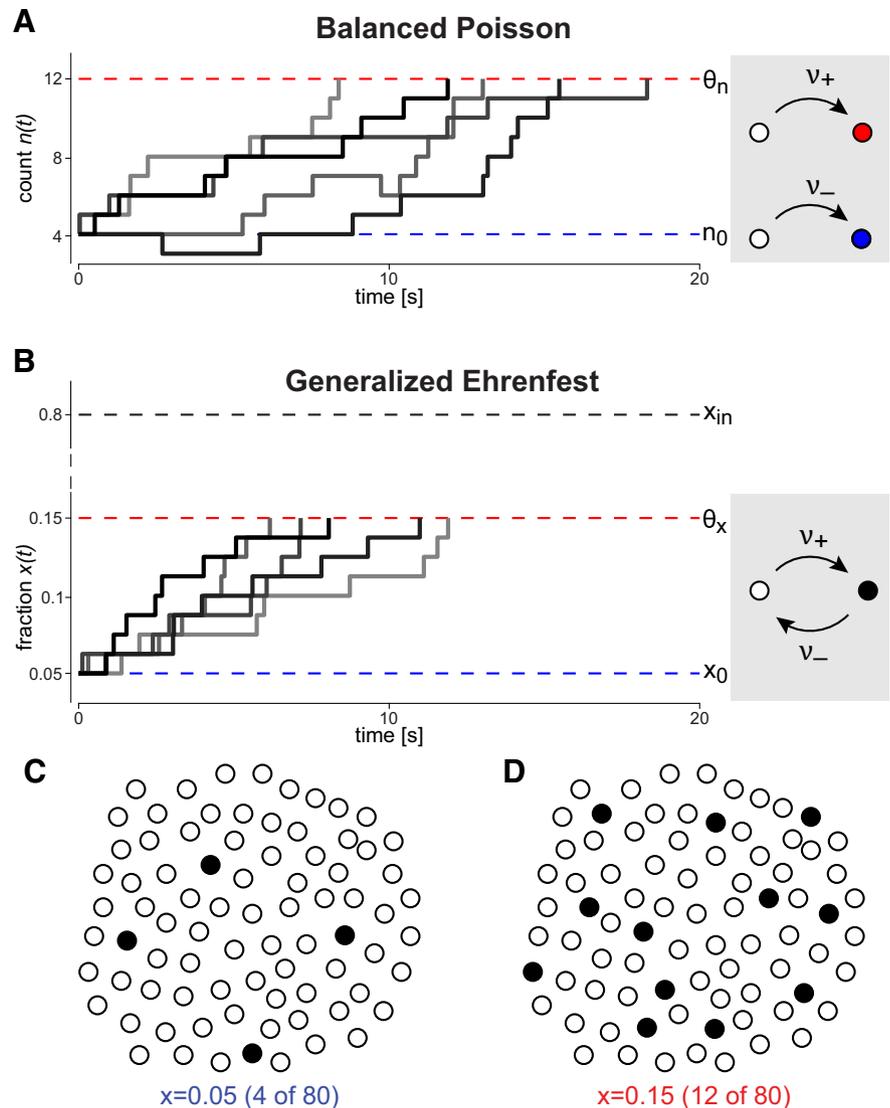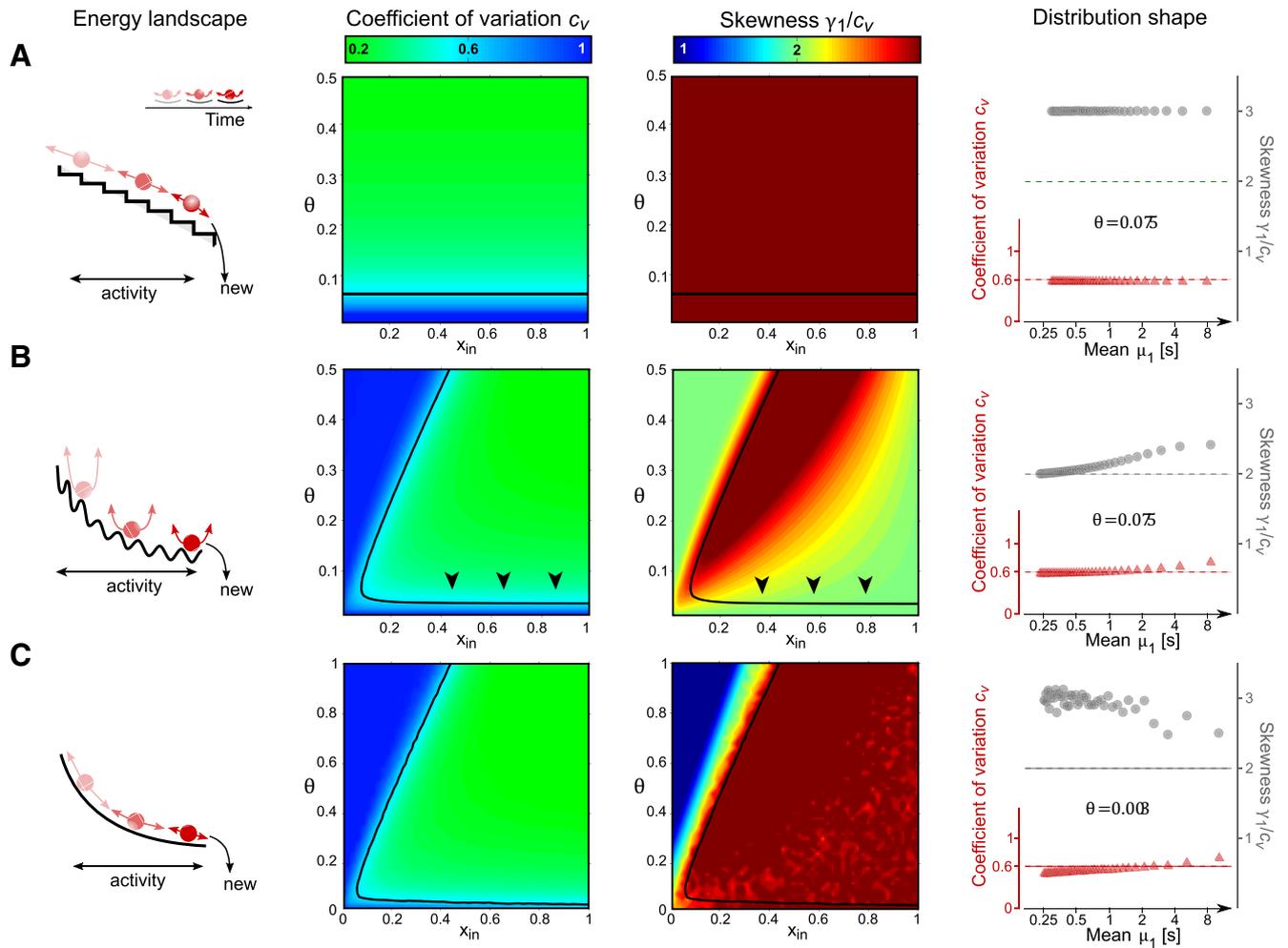

Figure 5. Discrete stochastic processes, including the proposed scenario for reproducing the scaling property of multistable perception. **A**, BP process counting independent increments and decrements with rates $\nu_E = 0.64$ Hz and $\nu_I = 16$ Hz, respectively. Accumulation begins at $n_0 = 4$ and ends at $\theta_n = 12$. Five realizations are shown (different shading). **B**, GE process of fractional activity $x(t)$ in a population of $N = 80$ units. Units are bistable and activate or inactivate spontaneously with rates $\nu_+ = 0.008$ Hz and $\nu_- = 0.002$ Hz, respectively. Accumulation begins at $x_0 = 0.05$ and ends at $\theta_x = 0.15$. Activity remains well below the steady-state activity level $x_{in} = \nu_+/(\nu_+ + \nu_-)$. **C**, At the start of an accumulation, only few units are active (fractional activation $x = 0.05$, 4 of 80). **D**, The accumulation ends when the fractional activation reaches threshold ($x = 0.15$, 12 of 80).

Figure 6. Processes with discrete stochastic events: normalized moments of FPT density as functions of input x_{in} and threshold θ. Leftmost column, Energy landscape and evolution over time (schematic). Middle left and right columns, Coefficient of variation c_v and skewness γ_1 (in units of c_v). Black curves represent $c_v = 0.6$. Black arrows indicate distribution shape of multistable perception ($c_v = 0.6$ and $\gamma_1 = 2c_v$). Rightmost column, Distribution shape as function of mean μ_1 (format as in Fig. 1B, C). For constant θ (value given by inset), μ_1 decreases as x_{in} increases from 0 to 1. **A**, Stochastic increments and decrements with input-dependent rates (homogeneous Poisson process). **B**, Stochastic increments and decrements generated by a finite population of discrete units (GE process). **C**, Limit of ever smaller increments and decrements generated by ever more units (continuous limit of GE process).

different (random) times and, superimposed, the progressive increase of the mean firing rate over all assemblies. The trajectory of collective activity is more predictable than are individual activations. If the threshold for collective activity is low, so that the entire accumulation proceeds far from steady-state, the FPTs of collective activity will reproduce the scaling property of multistable perception (compare Fig. 5B–D). Figure 8B shows a slightly different situation, with an initial steady-state fraction of $x_0 = 0.2$ active assemblies. In this case, individual assemblies both activate and (less often) inactivate at different (random) times. Once again, the progressive increase of collective activity can reproduce the scaling property of multistable perception.

For comparison, we also illustrate an alternative scenario that accumulates collective activity in a comparable way but cannot reproduce a scaling property (Fig. 8C). For this purpose, we simulated assemblies of weakly coupled spiking neurons, which do not express a bistable dynamics. Instead, the activity of each assembly fluctuates endogenously (finite-size noise) about a steady-state that is set by external input. To obtain a stochastic accumulation, we gradually increased external input, such as to raise the steady-state from low to high activity. The spike raster shows individual assemblies becoming progressively more active and, superimposed, the collective activity from all assemblies.

By construction, the alternative scenarios illustrated by Figure 8A, C are indistinguishable in terms of the time-varying average spike rate $r(t)$. (This is true for arbitrary subsets of neurons.) Nevertheless, even at the single-cell level, a bistable dynamics of individual assemblies is detectable in terms of the variability of firing rates and the distribution of interspike intervals. These telltale features are particularly evident compared with inhomogeneous Poisson spikes with identical average rate $r(t)$. Specifically, the SD of the instantaneous firing rate increases significantly during the rising phase of the accumulation (Fig. 8D) and the interspike interval distribution emphasizes extremes (short and long intervals) at the expense of the middle (Fig. 8E).

Discussion

We have examined the possibility that the highly consistent statistics of multistable phenomena may reflect the qualitative nature of perceptual representations. The macroscopic laws of physical systems often depend more on the qualitative nature than on the details of their microscopic constituents (Haken, 1975), and this may hold also for large-scale brain activity (Breakspear and Jirsa, 2007).
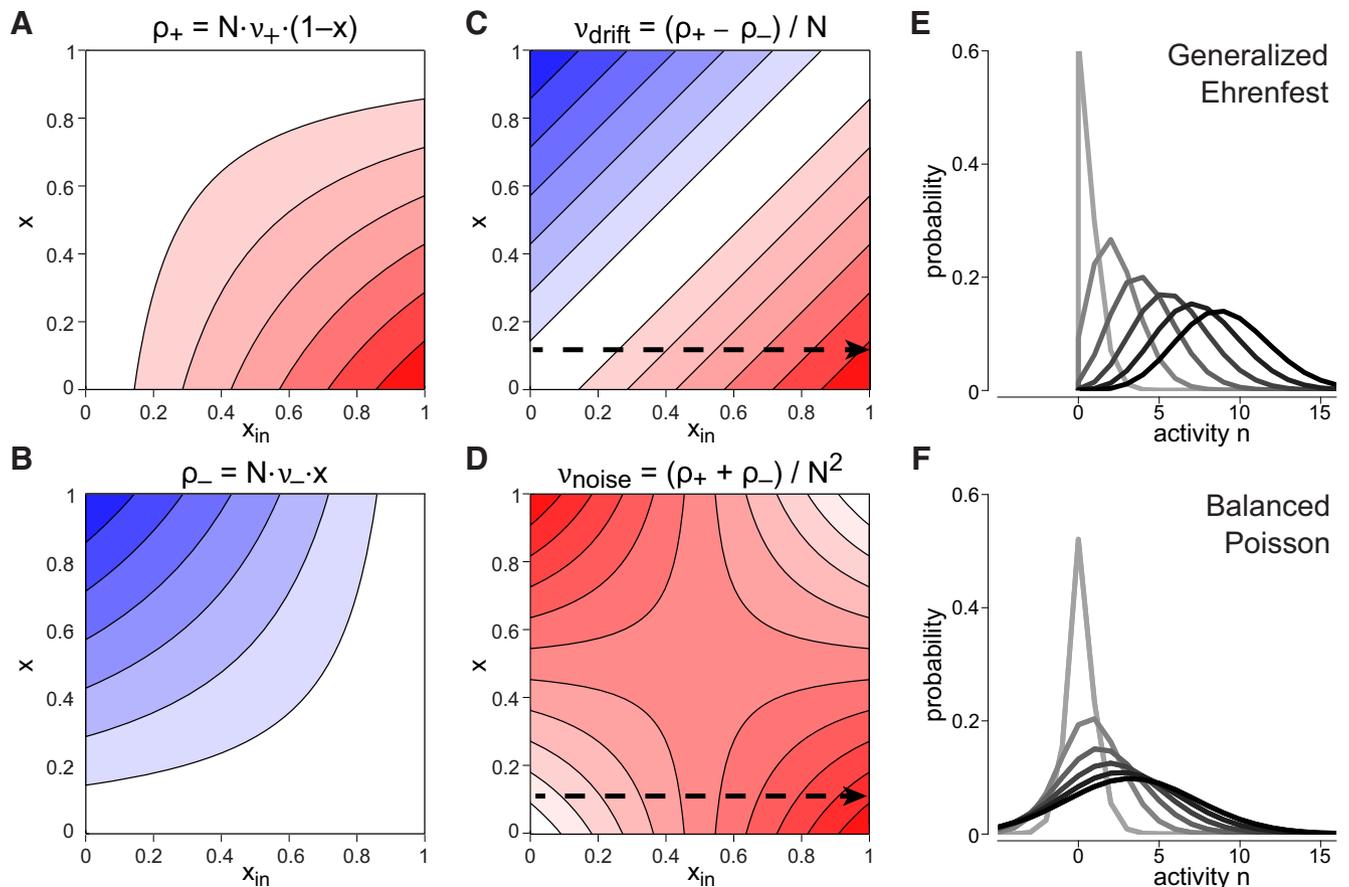
Figure 7. Origin of scaling property in a GE process. **A**, Increment rate ρ_+ (positive; red) increases with x_{in} and decreases with x. **B**, The same holds for the decrement rate ρ_- (negative; blue). **C**, Accumulation ν_{drift} varies with the difference $(\rho_+ - \rho_-)$. **D**, Dispersion ν_{noise} varies with the sum $(\rho_+ + \rho_-)$. For $x \ll 0.5$, both ν_{noise} and ν_{drift} change proportionally with x_{in} (dashed arrows), ensuring that normalized moments remain constant. **E**, Distribution of activity $x(t)$ at different t, with darker hues representing later times. Asymmetric dispersion reflects the constraint $x(t) \geq 0$. **F**, Same as **E** for BP process with identical event rates ($\nu_E = N\nu_+$, $\nu_I = N\nu_-$). Here, dispersion is symmetric, allowing low activity (including negative counts) to remain probable.

Current models of multistability

Current models of multistable perception seek to capture the interplay between stabilizing and destabilizing factors thought to underlie perceptual reversals. First, competition between alternative appearances stabilizes whichever currently dominates (Blake et al., 1990; Alais et al., 2010). Second, neural adaptation progressively weakens the dominant representation, shortening its persistence (Wolfe, 1984; Petersik, 2002; Kang and Blake, 2010). Third, neural noise initiates stochastic perceptual reversals (Hollins, 1980; Brascamp et al., 2006; Kim et al., 2006; Sterzer and Rees, 2008). Typically, such models seek to account for the occurrence of reversals, the rate of reversals, and its dependence on stimulus strength (Laing and Chow, 2002; Wilson, 2003; Moldakarimov et al., 2005; Moreno-Bote et al., 2007; Noest et al., 2007; Shpiro et al., 2007; Wilson, 2007; Shpiro et al., 2009; Seely and Chow, 2011; Pastukhov et al., 2013). Because these models combine a fast, stochastic process (noisy competition for dominance) with a slow, deterministic process (adaptation), they can be idealized as a fast escape with slowly adapting threshold.

The reasons why current models fail to reproduce the experimentally observed statistics are straightforward. First, they assume that the slow process (adaptation, corresponding to our accumulation rate ν_{drift}) changes with sensory input, whereas the fast process does not (noise, corresponding to our dispersion rate ν_{noise}). To obtain a scaling property, both adaptation and noise would have to vary with sensory input s (see Materials and

Methods). Second, current models assume symmetric dispersion, which necessarily yields heavier-tailed FPT distributions with $\gamma_1 \approx 3c_v$ (see Figs. 3A,B, 4A,B, 6C, 7E).

It is true that, for a given input strength, the experimentally observed skewness ($\gamma_1 \approx 2c_v$) may be obtained by precisely matching adaptation rate and noise amplitude (Pastukhov et al., 2013) (Fig. 3C, middle panels, arrows). However, to obtain a scaling property, both adaptation rate and noise amplitude would have to be matched anew for each input strength.

A finite population of bistable units accumulating activity far from equilibrium

To explain the observed statistics of multistable perception, we propose to replace the slow, deterministic process of current models (adaptation) with a discrete random walk (GE process). This is consistent with a suggestion of van Ee (2009), who has argued previously that reversal timing reflects a slow random walk rather than a deterministic relaxation, on the grounds that reversal timing is not entirely "memoryless" (Pastukhov and Braun, 2011). We do not see a need to change the fast, stochastic processes of current models, which provide the mechanism for reversals, and which in any case are beyond the scope of the present work.

Specifically, we propose that the neural representations underlying alternative perceptual appearances exhibit three features: (1) granularity in the sense of comprising a finite number of discrete
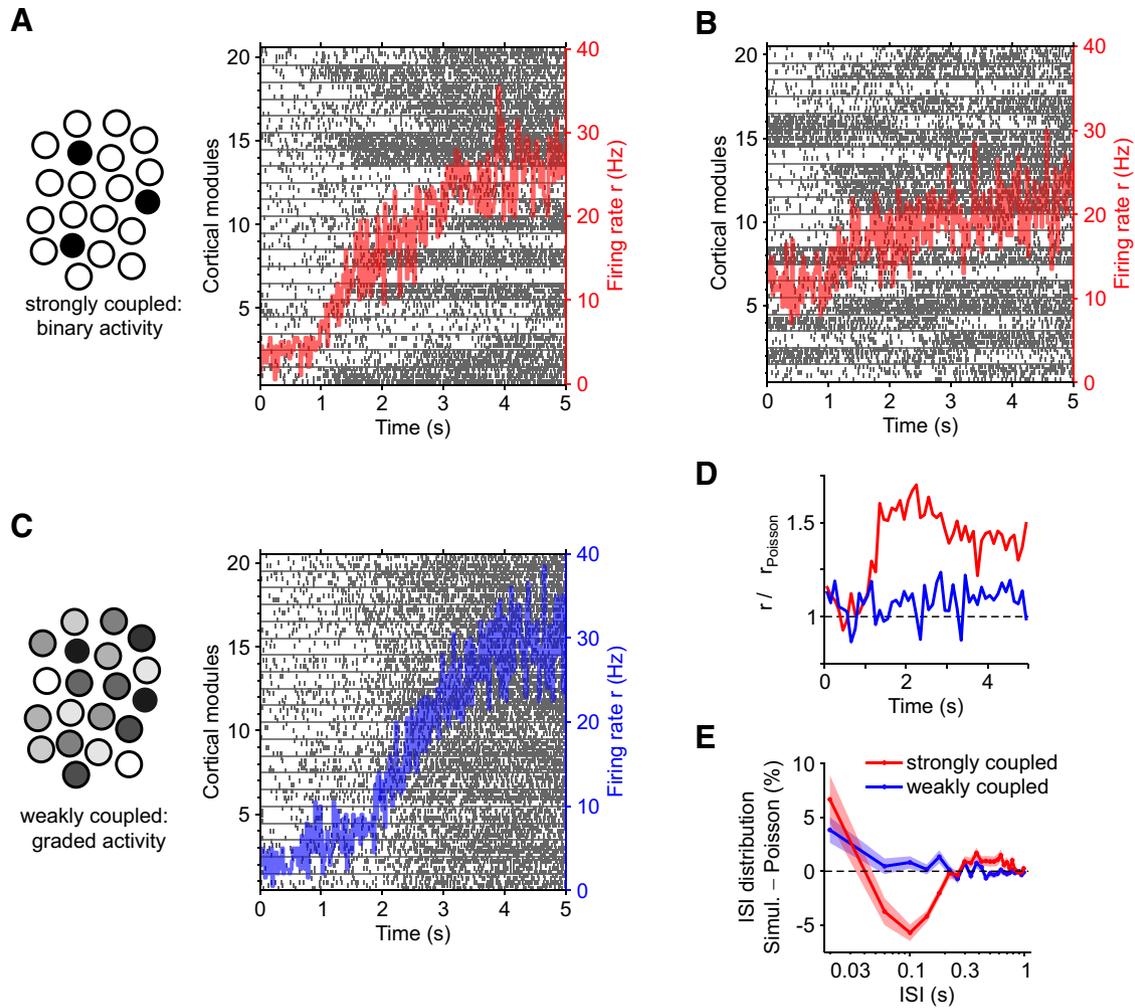
Figure 8. Stochastic accumulation of collective activity by modular assemblies of spiking neurons. ***A***, Strongly coupled assemblies expressing bistable attractor dynamics. Spike raster of five representative neurons per assembly (left ordinate) and collective activity $r(t)$ of all assemblies (red trace, right ordinate). Assemblies transition spontaneously and abruptly from inactive to active states. ***B***, Same as ***A***, but with some assemblies being active initially. Assemblies transition spontaneously to active states and (less often) to inactive states. ***C***, Weakly coupled assemblies without bistable dynamics. The activity of each assembly fluctuates about a steady-state level, which rises progressively due to external input. Collective activity $r(t)$ is comparable to ***A*** (blue trace, right ordinate). ***D***, SD Δr of mean firing rate $r(t)$ in ***A*** and ***C*** (red and blue trace, respectively), computed in 100 *m*s sliding windows relative to the SD $\Delta r_{Poisson}$ of inhomogeneous Poisson processes reproducing the observed $r(t)$. The SD may be computed either over neurons or over trials. ***E***, Distribution of interspike intervals (ISI) in ***A*** and ***C*** (red and blue trace, respectively), relative to the ISI distribution of inhomogeneous Poisson processes. Colored shading represents the SD of ISI density.

units; (2) local bistability in the sense that individual units transition spontaneously and independently between inactive and active states; and (3) far-from-equilibrium operation in the sense that the entire stochastic accumulation of discrete activity, from initial state to threshold, proceeds far from the steady-state.

Accumulating activity with bistable units ensures a scaling property because both accumulation rate ν_{drift} and dispersion rate ν_{noise} change proportionally with sensory input. Accumulating activity over a small range (from initial to threshold level) ensures the high variability of FPT that is experimentally observed. And accumulating activity far from equilibrium, where activity disperses asymmetrically, ensures the shorter-tailed FPT distributions that are observed (Figs. 3C, 7F). Two previous attempts to model multistable perception with populations of bistable units (Taylor and Aldridge, 1974; Gigante et al., 2009) failed to reproduce these statistics (De Marco et al., 1977) because neither operated far enough from equilibrium.

Collective dynamics of cortical columns or clusters

To what extent is the scenario outlined above consistent with, or even supported by, the available evidence about the dynamics of neocortical activity?

The postulated units idealize the well-known dynamics of attractor assemblies (Amit, 1995). A neuronal assembly with balanced excitation and inhibition may express two distinct steady-states, a low activity (inactive state) and a high activity state (active state), provided that recurrent excitation amplifies inputs nonlinearly (Amit and Brunel, 1997; Wang, 2002; Major and Tank, 2004). If this is the case, collective activity is bistable and, driven by endogeneous fluctuations (finite-size noise), transitions spontaneously between the two steady-states. External inputs alter the relative stability of inactive and active states and therefore modulate the transition rates. In sum, balanced attractor assemblies exhibit all postulated properties: local bistability, noise-driven transitions, and input-dependent rates.

The crucial ingredient of attractor dynamics, recurrent excitatory connectivity, is well established for cortical columns and clusters of columns. In the superficial layers of primate neocortex, neurons are organized into radial columns with similar response properties (Snodderly and Gur, 1995; Mountcastle, 1997; Maier et al., 2010) and preferential recurrent connectivity (Douglas and Martin, 2004). Nearby columns with similar properties are linked by "patchy" horizontal connections into locally dispersed clusters (Bosking et al., 1997; Lund et al., 2003; Tanigawa et al., 2005; Muir and Douglas, 2011). The preferential excitatory connectivity within clusters of columns shapes both the response properties of individual neurons (Angelucci and Bressloff, 2006) and the "patchy" patterns of activity evoked by stimulation (Arieli et al., 1996; Bosking et al., 1997; Tsodyks et al., 1999).

Neocortical activity varies spontaneously even in the absence of external stimulation and to a degree comparable to that evoked by stimulation (Arieli et al., 1996; Destexhe et al., 2003). Interestingly, the spatiotemporal structure of spontaneous activity often corresponds closely to activity patterns evoked by sensory stimulation (Kenet et al., 2003; Petersen et al., 2003; MacLean et al., 2005; Luczak et al., 2007; Sakata and Harris, 2009; Harris et al., 2011). On the basis of these observations, it has been suggested that cortex expresses only a limited vocabulary of activity patterns and that this vocabulary is explored more extensively by spontaneous activity than by evoked activity (Luczak et al., 2009; Harris et al., 2011).

A slow exploration of a limited vocabulary of activity states is exactly the dynamics expressed by interacting attractor assemblies (Durstewitz and Deco, 2008; Litwin-Kumar and Doiron, 2012; Wang, 2012). If excitatory connectivity is not uniform but forms clusters of preferentially connected neurons, then these clusters may express a slow dynamics (<10 Hz), transiently increasing or decreasing their collective activity (Litwin-Kumar and Doiron, 2012). Despite the slow correlation, the timing of individual spikes in a cluster remains asynchronous (Renart et al., 2010). External stimulation biases the slow dynamics and, for stimulated clusters, selectively extends the periods of increased activity.

Whether or not cortical columns or clusters do indeed express bona fide attractor dynamics remains an open question. Evidence from anesthetized rodents and from *in vitro* preparations shows that, at a minimum, cortical networks may enter into transient states of enhanced excitability, which are characterized by a consistently depolarized membrane potential, asynchronous spiking activity, and nonlinear amplification of inputs (Haider and McCormick, 2009; Harris et al., 2011). These states are maintained by a barrage of excitatory and inhibitory synaptic potentials from recurrent connectivity.

Recent evidence from awake rodents and primates demonstrates that many (not all) cortical neurons are strongly coupled to local population activity (Okun et al., 2015), consistent with a local attractor assembly. Sophisticated statistical approaches have recovered evidence for sudden transitions between discrete activity states in sensory, motor, and parietal cortices of behaving animals (Abeles et al., 1995; Jones et al., 2007; Latimer et al., 2015). Even more intriguingly, laminar recording in extrastriate visual cortex of awake primates revealed columnar activity to be bimodal, transitioning between inactive and active states "nearly simultaneously throughout the cortical dept" (Engel et al., 2015). Transition times were random (Poisson statistics), and transition rates were modulated by stimulus and task conditions, exactly as expected from locally bistable assemblies (Engel et al., 2015). In premotor cortex of awake primates, local multiunit activity was found to be bimodal at 57 of 107 recorded locations (Mattia et al., 2013). Additionally, the differential power spectrum of multiunit

activity was found to be consistent with bistable attractor dynamics at a subset of recorded locations.

Conclusions and outlook

We have shown that the characteristic statistics of multistable perception would be reproduced by perceptual representations that comprise populations of attractor assemblies, such as cortical columns or clusters of columns, each expressing a bistable local dynamics, and that accumulate activity in a narrow range far from equilibrium. This idealized scenario may be relaxed without compromising the FPT statistics: the population may be heterogeneous in that assemblies may display different transition rates, may encode different aspects of the stimulus, and may be situated at different levels of a cortical hierarchy, provided all assemblies contribute additively to collective activity.

The perceptual representation postulated here is well suited to support reliable probabilistic inference by stochastic sampling (Hinton and Sejnowski, 1986; Buesing et al., 2011). Specifically, Langevin sampling by a modified Boltzmann machine (Geman and Geman, 1984; Hoyer and Hyvärinen, 2003; Ma et al., 2006; Sundareswara and Schrater, 2008; Churchland et al., 2011; Gershman et al., 2012) offers a functional rationale as to why competing perceptual representations should operate far from equilibrium and at a low level of fractional activation: in this regimen, microtrajectories involving only a few transitions may be expected to acquire maximal influence (maximum caliber) on the competition outcome (Ghosh et al., 2006; Pressé et al., 2013), permitting inferences to be drawn with optimal reliability.

References

Abeles M, Bergman H, Gat I, Meilijson I, Seidemann E, Tishby N, Vaadia E (1995) Cortical activity flips among quasi-stationary states. Proc Natl Acad Sci U S A 92:8616–8620. CrossRef Medline

Adelson EH, Movshon JA (1982) Phenomenal coherence of moving visual patterns. Nature 300:523–525. CrossRef Medline

Alais D, Cass J, O'Shea RP, Blake R (2010) Visual sensitivity underlying changes in visual consciousness. Curr Biol 20:1362–1367. CrossRef Medline

Amit D (1995) The Hebbian paradigm reintegrated: local reverberations as internal representations. Behav Brain Sci 18:617–625. CrossRef

Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb Cortex 7:237–252. CrossRef Medline

Angelucci A, Bressloff PC (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extraclassical receptive field surround of primate V1 neurons. Prog Brain Res 154:93–120. CrossRef Medline

Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273:1868–1871. CrossRef Medline

Blake R, Logothetis N (2002) Visual competition. Nat Rev Neurosci 3: 13–21. CrossRef Medline

Blake RR, Fox R, McIntyre C (1971) Stochastic properties of stabilized-image binocular rivalry alternations. J Exp Psychol 88:327–332. CrossRef Medline

Blake R, Westendorf D, Fox R (1990) Temporal perturbations of binocular rivalry. Percept Psychophys 48:593–602. CrossRef Medline

Bonneh YS, Cooperman A, Sagi D (2001) Motion-induced blindness in normal observers. Nature 411:798–801. CrossRef Medline

Borsellino A, De Marco A, Allazetta A, Rinesi S, Bartolini B (1972) Reversal time distribution in the perception of visual ambiguous stimuli. Kybernetik 10:139–144. CrossRef Medline

Bosking WH, Zhang Y, Schofield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J Neurosci 17:2112–2127. Medline

Brascamp JW, van Ee R, Pestman WR, van den Berg AV (2005) Distributions of alternation rates in various forms of bistable perception. J Vis 5:287–298. CrossRef Medline

Brascamp JW, van Ee R, Noest AJ, Jacobs RH, van den Berg AV (2006) The time course of binocular rivalry reveals a fundamental role of noise. J Vis 6:1244–1256. CrossRef Medline

Braun J, Mattia M (2010) Attractors and noise: twin drivers of decisions and multistability. Neuroimage 52:740–751. CrossRef Medline

Breakspear M, Jirsa V (2007) Neuronal dynamics and brain connectivity. In: Handbook of brain connectivity (Jirsa V, McIntosh A, eds), pp 3–64. New York: Springer.

Bregman AS (1994) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: Massachusetts Institute of Technology.

Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci 8:183–208. CrossRef Medline

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

Bundesen C, Vangkilde S, Petersen A (2015) Recent developments in a computational theory of visual attention (TVA). Vision Res 116:210–218. CrossRef Medline

Campbell FW, Howell ER (1972) Monocular alternation: a method for the investigation of pattern vision. J Physiol 225:19P–21P. Medline

Cao R, Braun J, Mattia M (2014) Stochastic accumulation by cortical columns may explain the scalar property of multi-stable perception. Phys Rev Lett 113:098103. CrossRef Medline

Carpenter RH (2012) Analysing the detail of saccadic reaction time distribution. Biocybern Biomed Eng 32:49–63. CrossRef

Churchland AK, Kiani R, Chaudhuri R, Wang XJ, Pouget A, Shadlen MN (2011) Variance as a signature of neural computations during decision making. Neuron 69:818–831. CrossRef Medline

Cox DR, Miller HD (1972) The theory of stochastic processes. London: Chapman and Hall.

Cox JC, Ingersoll JE, Ross SA (1985) A theory of the term structure of interest rates. Econometrica 53:385–407. CrossRef

De Marco A, Penengo P, Trabucco A (1977) Stochastic models and fluctuations in reversal time of ambiguous figures. Perception 6:645–656. CrossRef Medline

Destexhe A, Rudolph M, Paré D (2003) The high-conductance state of neocortical neurons in vivo. Nat Rev Neurosci 4:739–751. CrossRef Medline

Douglas RJ, Martin KA (2004) Neuronal circuits of the neocortex. Annu Rev Neurosci 27:419–451. CrossRef Medline

Durstewitz D, Deco G (2008) Computational significance of transient dynamics in cortical networks. Eur J Neurosci 27:217–227. CrossRef Medline

Emch G, Liu C (2002) The logic of thermostatistical physics. New York: Springer.

Engel T, Steinmetz N, Moore T, Boahen K (2015) Attentional modulation of cortical state dynamics and its contribution to spiking variability in area V4. In: Society for Neuroscience Abstracts (372.07). Chicago: Society for Neuroscience.

Fox R, Herrmann J (1967) Stochastic properties of binocular rivalry alternations. Percept Psychophys 2:432–436. CrossRef

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741. Medline

Gershman SJ, Vul E, Tenenbaum JB (2012) Multistability and perceptual inference. Neural Comput 24:1–24. CrossRef Medline

Gigante G, Mattia M, Braun J, Del Giudice P (2009) Bistable perception modeled as competing stochastic integrations at two levels. PLoS Comput Biol 5:e1000430. CrossRef Medline

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574. CrossRef Medline

Ghosh K, Dill KA, Inamdar MM, Seitaridou E, Phillips R (2006) Teaching the principles of statistical dynamics. Am J Phys 74:123–133. CrossRef Medline

Haider B, McCormick DA (2009) Rapid neocortical dynamics: cellular and network mechanisms. Neuron 62:171–189. CrossRef Medline

Haken H (1975) Cooperative phenomena in systems far from thermal equilibrium and in nonphysical systems. Rev Modern Phys 47:67–121. CrossRef

Harris KD, Bartho P, Chadderton P, Curto C, de la Rocha J, Hollender L, Itskov V, Luczak A, Marguet SL, Renart A, Sakata S (2011) How do neurons work together? Lessons from auditory cortex. Hear Res 271: 37–53. CrossRef Medline

Hinton GE, Sejnowski TJ (1986) Learning and relearning in Boltszmann machines. In: Parallel distributed processing: explorations in the microstructure of cognition, Vol 1 (Rumelhart G, McClelland J, eds). Foundations. Cambridge, MA: Massachusetts Institute of Technology.

Hohwy J, Roepstorff A, Friston K (2008) Predictive coding explains binocular rivalry: an epistemological review. Cognition 108:687–701. CrossRef Medline

Hollins M (1980) The effect of contrast on the completeness of binocular rivalry suppression. Percept Psychophys 27:550–556. CrossRef Medline

Hoyer PO, Hyvärinen A (2003) Interpreting neural response variability as Monte Carlo sampling of the posterior. Adv Neur Inf Proc Sys 15:293–300.

Inoue J, Sato S, Ricciardi LM (1995) On the parameter estimation for diffusion models of single neuron's activities. Biol Cybern 73:209–221. CrossRef Medline

Jones LM, Fontanini A, Sadacca BF, Miller P, Katz DB (2007) Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. Proc Natl Acad Sci U S A 104:18772–18777. CrossRef Medline

Kang MS, Blake R (2010) What causes alternations in dominance during binocular rivalry? Atten Percept Psychophys 72:179–186. CrossRef Medline

Karlin S, McGregor JL (1965) Ehrenfest Urn Models. J Appl Prob 2: 352–376. CrossRef

Kenet T, Bibitchkov D, Tsodyks M, Grinvald A, Arieli A (2003) Spontaneously emerging cortical representations of visual attributes. Nature 425: 954–956. CrossRef Medline

Kim YJ, Grabowecky M, Suzuki S (2006) Stochastic resonance in binocular rivalry. Vision Res 46:392–406. CrossRef Medline

Knapen T, Brascamp J, Pearson J, van Ee R, Blake R (2011) The role of frontal and parietal brain areas in bistable perception. J Neurosci 31: 10293–10301. CrossRef Medline

Laing CR, Chow CC (2002) A spiking neuron model for binocular rivalry. J Comput Neurosci 12:39–53. CrossRef Medline

Latimer KW, Yates JL, Meister ML, Huk AC, Pillow JW (2015) Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science 349:184–187. CrossRef Medline

Leopold DA, Logothetis NK (1999) Multistable phenomena: changing views in perception. Trends Cogn Sci 3:254–264. CrossRef Medline

Levelt WJ (1967) Note on the distributions of dominance times in binocular rivalry. Br J Psychol 58:143–145. CrossRef Medline

Litwin-Kumar A, Doiron B (2012) Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat Neurosci 15: 1498–1505. CrossRef Medline

Luczak A, Barthó P, Marguet SL, Buzsáki G, Harris KD (2007) Sequential structure of neocortical spontaneous activity in vivo. Proc Natl Acad Sci U S A 104:347–352. CrossRef Medline

Luczak A, Barthó P, Harris KD (2009) Spontaneous events outline the realm of possible sensory responses in neocortical populations. Neuron 62: 413–425. CrossRef Medline

Lund JS, Angelucci A, Bressloff PC (2003) Anatomical substrates for functional columns in macaque monkey primary visual cortex. Cereb Cortex 13:15–24. CrossRef Medline

Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nat Neurosci 9:1432–1438. CrossRef Medline

MacLean JN, Watson BO, Aaron GB, Yuste R (2005) Internal dynamics determine the cortical response to thalamic stimulation. Neuron 48: 811–823. CrossRef Medline

Maier A, Adams GK, Aura C, Leopold DA (2010) Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. Front Syst Neurosci 4:31. CrossRef Medline

Major G, Tank D (2004) Persistent neural activity: prevalence and mechanisms. Curr Opin Neurobiol 14:675–684. CrossRef Medline

Mattia M, Pani P, Mirabella G, Costa S, Del Giudice P, Ferraina S (2013) Heterogeneous attractor cell assemblies for motor planning in premotor cortex. J Neurosci 33:11155–11168. CrossRef Medline

Meng M, Tong F (2004) Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. J Vis 4:539–551. CrossRef Medline

Moldakarimov S, Rollenhagen JE, Olson CR, Chow CC (2005) Competitive dynamics in cortical responses to visual stimuli. J Neurophysiol 94: 3388–3396. CrossRef Medline

Molini A, Talkner P, Katul G (2011) First passage time statistics of Brownian

motion with purely time dependent drift and diffusion. Physica A 390: 1841–1852. CrossRef

Moreno-Bote R, Rinzel J, Rubin N (2007) Noise-induced alternations in an attractor network model of perceptual bistability. J Neurophysiol 98: 1125–1139. CrossRef Medline

Moreno-Bote R, Shpiro A, Rinzel J, Rubin N (2010) Alternation rate in perceptual bistability is maximal at and symmetric around equidominance. J Vis 10:1. CrossRef Medline

Mountcastle VB (1997) The columnar organization of the neocortex. Brain 120:701–722. CrossRef Medline

Muir DR, Douglas RJ (2011) From neural arbors to daisies. Cereb Cortex 21:1118–1133. CrossRef Medline

Murata T, Matsui N, Miyauchi S, Kakita Y, Yanagida T (2003) Discrete stochastic process underlying perceptual rivalry. Neuroreport 14: 1347–1352. CrossRef Medline

Nawrot M, Blake R (1989) Neural integration of information specifying structure from stereopsis and motion. Science 244:716–718. CrossRef Medline

Noest AJ, van Ee R, Nijs MM, van Wezel RJA (2007) Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. J Vis 7:10. CrossRef Medline

Okun M, Steinmetz NA, Cossell L, Iacaruso MF, Ko H, Barthó P, Moore T, Hofer SB, Mrsic-Flogel TD, Carandini M, Harris KD (2015) Diverse coupling of neurons to populations in sensory cortex. Nature 512: 511–515. CrossRef Medline

Pastukhov A, Braun J (2007) Perceptual reversals need no prompting by attention. J Vis 7:5.1–17. CrossRef Medline

Pastukhov A, Braun J (2011) Cumulative history quantifies the role of neural adaptation in multistable perception. J Vis 11:12. CrossRef Medline

Pastukhov A, García-Rodriguez PE, Haenicke J, Guillamon A, Deco G, Braun J (2013) Cumulative history quantifies the role of neural adaptation in multistable perception. Front Comput Neurosci 7:17. Medline

Pearson B, Raskevicius J, Bays PM, Pertzov Y, Husain M (2014) Working memory retrieval as a decision process. J Vis 14:1–15. CrossRef Medline

Pearson J, Clifford CW, Tong F (2008) The functional impact of mental imagery on conscious perception. Curr Biol 18:982–986. CrossRef Medline

Petersen CC, Hahn TT, Mehta M, Grinvald A, Sakmann B (2003) Interaction of sensory responses with spontaneous depolarization in layer 2/3 barrel cortex. Proc Natl Acad Sci U S A 100:13638–13643. CrossRef Medline

Petersik JT (2002) Buildup and decay of a three-dimensional rotational aftereffect obtained with a three-dimensional figure. Perception 31: 825–836. CrossRef Medline

Pressé S, Ghosh K, Lee J, Dill KA (2013) Principles of maximum entropy and maximum caliber in statistical physics. Rev Modern Phys 85: 1115–1141. CrossRef

Ratcliff R, Smith PL (2004) A comparison of sequential sampling models for two-choice reaction time. Psychol Rev 111:333–367. CrossRef Medline

Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD (2010) The asynchronous state in cortical circuits. Science 327:587–590. CrossRef Medline

Resulaj A, Kiani R, Wolpert DM, Shadlen MN (2009) Changes of mind in decision-making. Nature 461:263–266. CrossRef Medline

Risken H (1984) Fokker-Planck equation. In: The Fokker-Planck equation, Vol 18. Springer series in synergetics, pp 63–95. Berlin: Springer.

Sakata S, Harris KD (2009) Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. Neuron 64: 404–418. CrossRef Medline

Schall JD (2001) Neural basis of deciding, choosing and acting. Nat Rev Neurosci 2:33–42. CrossRef Medline

Seely J, Chow CC (2011) Role of mutual inhibition in binocular rivalry. J Neurophysiol 106:2136–2150. CrossRef Medline

Shadlen MN, Kiani R (2013) Decision making as a window to cognition. Neuron 80:791–806. CrossRef Medline

Shpiro A, Curtu R, Rinzel J, Rubin N (2007) Dynamical characteristics common to neuronal competition models. J Neurophysiol 97:462–473. CrossRef Medline

Shpiro A, Moreno-Bote R, Rubin N, Rinzel J (2009) Balance between noise and adaptation in competition models of perceptual bistability. J Comput Neurosci 37–54.

Smith PL (2000) Stochastic dynamic models of response time and accuracy: a foundational primer. J Math Psychol 44:408–463. CrossRef Medline

Smith PL, Ratcliff R (2009) An integrated theory of attention and decision making in visual signal detection. Psychol Rev 116:283–317. CrossRef Medline

Smith PL, Ratcliff R, McKoon G (2014) The diffusion model is not a deterministic growth model: comment on Jones and Dzhafarov (2014). Psychol Rev 121:679–688. CrossRef Medline

Snodderly DM, Gur M (1995) Organization of striate cortex of alert, trained monkeys (*Macaca fascicularis*): ongoing activity, stimulus selectivity, and widths of receptive field activating regions. J Neurophysiol 74:2100–2125. Medline

Sperling G, Dosher BA (1994) Depth from motion. In: Early vision and beyond (Papathomas TV, Charles Chubb AG, Kowler E, eds), pp 133–142. Cambridge, MA: Massachusetts Institute of Technology.

Sterzer P, Rees G (2008) A neural basis for percept stabilization in binocular rivalry. J Cogn Neurosci 20:389–399. CrossRef Medline

Sterzer P, Kleinschmidt A, Rees G (2009) The neural bases of multistable perception. Trends Cogn Sci 13:310–318. CrossRef Medline

Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci 6:363–375. CrossRef Medline

Sundareswara R, Schrater PR (2008) Perceptual multistability predicted by search model for Bayesian decisions. J Vis 8:1–19. CrossRef Medline

Tanigawa H, Wang Q, Fujita I (2005) Organization of horizontal axons in the inferior temporal cortex and primary visual cortex of the macaque monkey. Cereb Cortex 15:1887–1899. CrossRef Medline

Taylor M, Aldridge K (1974) Stochastic processes in reversing figure perception. Percept Psychophys 16:9–25. CrossRef

Tsodyks M, Kenet T, Grinvald A, Arieli A (1999) Linking spontaneous activity of single cortical neurons and the underlying functional architecture. Science 286:1943–1946. CrossRef Medline

Tuckwell HC (2001/2008) Introduction to theoretical neurobiology, Vol 2. Nonlinear and stochastic theories. Cambridge, MA: Cambridge UP.

Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. Psychol Rev 108:550–592. CrossRef Medline

van Ee R (2009) Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: evidence from serial correlations in perceptual bistability. J Opt Soc Am A 26:2612–2622. CrossRef Medline

van Kampen NG (1981) Stochastic processes in physics and chemistry. Amsterdam: North-Holland Physics.

von Grünau M, Dubé S (1993) Ambiguous plaids: switching between coherence and transparency. Spatial Vision 7:199–211. Medline

Walker P (1975) Stochastic properties of binocular rivalry alternations. Percept Psychophys 18:467–473. CrossRef

Wallach H, O'Connell DN (1953) The kinetic depth effect. J Exp Psychol 45:205–217. CrossRef Medline

Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36:955–968. CrossRef Medline

Wang XJ (2012) Neural dynamics and circuit mechanisms of decision-making. Curr Opin Neurobiol 22:1039–1046. CrossRef Medline

Wilson HR (2003) Computational evidence for a rivalry hierarchy in vision. Proc Natl Acad Sci U S A 100:14499–14503. CrossRef Medline

Wilson HR (2007) Minimal physiological conditions for binocular rivalry and rivalry memory. Vision Res 47:2741–2750. CrossRef Medline

Winkler I, Denham S, Mill R, Bohm TM, Bendixen A (2012) Multistability in auditory stream segregation: a predictive coding view. Philos Trans R Soc Lond B Biol Sci 367:1001–1012. CrossRef Medline

Wolfe JM (1984) Reversing ocular dominance and suppression in a single flash. Vision Res 24:471–478. CrossRef Medline

Zhou YH, Gao JB, White KD, Merk I, Yao K (2004) Perceptual dominance time distributions in multistable visual perception. Biol Cybern 90: 256–263. CrossRef Medline