Critical and maximally informative encoding between neural populations in the retina

David B. Kastner[a,b,1], Stephen A. Baccus[c], and Tatyana O. Sharpee[b,d,2]

[a]Neuroscience Program and [c]Department of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305; [b]Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037; and [d]Center for Theoretical Biological Physics and Department of Physics, University of California, San Diego, La Jolla, CA 92093

Computation in the brain involves multiple types of neurons, yet the organizing principles for how these neurons work together remain unclear. Information theory has offered explanations for how different types of neurons can maximize the transmitted information by encoding different stimulus features. However, recent experiments indicate that separate neuronal types exist that encode the same filtered version of the stimulus, but then the different cell types signal the presence of that stimulus feature with different thresholds. Here we show that the emergence of these neuronal types can be quantitatively described by the theory of transitions between different phases of matter. The two key parameters that control the separation of neurons into subclasses are the mean and standard deviation (SD) of noise affecting neural responses. The average noise across the neural population plays the role of temperature in the classic theory of phase transitions, whereas the SD is equivalent to pressure or magnetic field, in the case of liquid–gas and magnetic transitions, respectively. Our results account for properties of two recently discovered types of salamander Off retinal ganglion cells, as well as the absence of multiple types of On cells. We further show that, across visual stimulus contrasts, retinal circuits continued to operate near the critical point whose quantitative characteristics matched those expected near a liquid–gas critical point and described by the nearest-neighbor Ising model in three dimensions. By operating near a critical point, neural circuits can maximize information transmission in a given environment while retaining the ability to quickly adapt to a new environment.

neural coding | information theory | phase transitions | scaling

Neural circuits use populations composed of multiple cell types to perform complex computations. Theoretical arguments, based upon the maximization of information transmitted about incoming stimuli, have proved successful in accounting for properties of single neurons (1–4) or populations of neurons encoding either one (5–9) or several different visual features (10–13). However, recent experiments in the retina have discovered types of neurons whose responses are triggered by the presence of the same visual feature in the stimulus but differ in the threshold value by which they detect and report the presence of that feature, where threshold is defined as 50% spiking probability in a binary neuron (Fig. 1A) (14). These neurons may be described by having the same linear properties (reflecting the match in their relevant stimulus features) but different nonlinear properties, such as different threshold values. Here we sought to develop a framework to explain the existence of such neuronal types as a function of the average noise across the neural population and differences in the noise between neural classes. We use the retina as a tractable system to study how neural responses in heterogeneous populations might be coordinated to efficiently encode complex sensory inputs.

In the salamander retina, populations of two types of Off cells encode nearly the same spatiotemporal visual feature, but separately tile the retina, indicating that they are distinct cell types (14). These two types of neurons also maintain different

thresholds across a range of contrasts, with differences in threshold between populations exceeding the variation in threshold within each type (Fig. S1). Specifically, fast Off sensitizing cells maintain a lower threshold, encoding weaker signals, than fast Off adapting cells, which encode stronger signals. Notably, all types of Off cells split into such adapting and sensitizing subtypes, whereas the On types do not. These specific differences in the neural encoding between populations provide a particularly convenient model in which to analyze the factors that cause distinct neural populations to arise.

Here we show that the splitting of a population into two cell types with different detection thresholds can be explained in terms of maximization of information transmission and that the decision to split or not is an abrupt transition point that depends critically on the noise level. We further show that the process of optimizing information transmission in neural populations facing the decision to transition between one and two populations is mathematically identical to optimization of free energy in physical systems undergoing transitions between different phases of matter. Although the signatures that neural circuits might be poised near a phase transition have been observed (15–19), the correspondence to phase transitions in physical systems (16) and the relevance to neuroscience have been debated (19). This correspondence between neural and physical systems allows us to infer that retinal populations adopt a specific position relative to a critical point to increase information transmission in one environment, yet still retain the ability to adjust quickly if the environment changes.

Results

A Model for Quantifying Information Transmission in a Population of Multiple Types of Neurons. In the retina, neurons of a given cell type tile visual space. The full quantification and derivation of

Significance

It is unknown what functional properties influence the number of cell types in the brain. Here we show how one can use a powerful framework from physics that describes the transitions between different phases of matter, such as between liquid and gas, to specify under what conditions it becomes optimal to split neural populations into new subtypes to maximize information transmission. These results outline a conceptual framework that spans both physical and biological systems and can be used to explain the emergence of different functional classes of neuronal types.
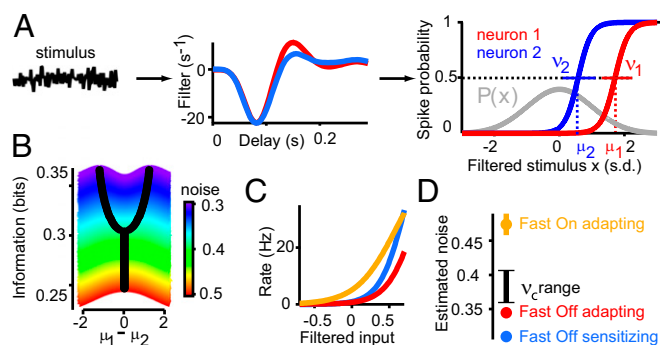
Fig. 1. Bifurcation between maximally informative solutions accounts for differences in homogeneity of On and Off ganglion cell populations. (*A*) Encoding model for neural responses. White noise stimuli (*Left*) are filtered by a convolution with a temporal filter (*Center*) that describes the time course of the preferred stimulus feature of a given neuron. These profiles are similar for sensitizing (blue) and adapting (red) cells and we neglect differences between them. (*Right*) The filtered stimulus is then passed through a static nonlinear function that maps the filtered stimulus onto a spiking probability that ranges between zero and one. The threshold μ is the value of the filtered stimulus x that yields 50% spiking probability, and the slope ν reflects noise that results in fractional spiking probabilities. (*B*) Information as a function of the threshold difference of the two response functions with the same slope (denoted by color). The rate constraint, $\langle r \rangle$, was the same for all curves shown. Black line shows the location of the maximal information for a given slope value. Bifurcation indicates the transition in the optimal solution from one to two populations. (*C*) On cells have shallower slopes than fast Off cells. Example nonlinearities for a fast Off sensitizing, a fast Off adapting, and an On cell are recorded simultaneously. Input values are normalized such that their SD is equal to 1. (*D*) Average slopes for fast Off sensitizing ($n = 95$), fast Off adapting ($n = 388$), and On ($n = 58$) cells. The critical point ν_c depends on the average rate $\langle r \rangle$ and is shown for the range of rates found in the data. For each $\langle r \rangle$, ν_c was determined by fitting a power law to the dependence of optimal threshold difference m on noise $\nu - \nu_c$, from B (Figs. S3C and S4A).

how and why interacting arrays of retinal neurons jointly encode the visual world remains computationally challenging. However, the information transmitted by multiple arrays of different types of neurons is expected to be proportional to the amount of information transmitted about a small local segment of visual space by neurons with receptive fields in that region (8, 10). Thus, we simplify the problem here by focusing on sets of cells located in a single region of space instead of having to account for a full spatial array of cells and on those aspects of visual stimuli that are encoded by a set of one fast Off adapting neuron and one fast Off sensitizing neuron. Because these neurons encode the same type of visual feature, we can consider the encoding of a one-dimensional signal, reflecting the visual stimulus filtered by this feature, which is then converted into spikes by a static nonlinear encoding function with a sigmoidal shape (Fig. 1*A*). Our goal was to find such nonlinear encoding functions that would maximize the accuracy with which this filtered stimulus is represented in the responses of the two neurons. We note that reducing two arrays of neurons to two representative neurons corresponds to the mean-field approximation in physics, which has proved successful in describing some of the most prominent aspects of emergent collective phenomena in physical systems (20).

Because the response of a spiking neuron in a small time interval can take only two values (zero or one spike), we modeled responses for each neuron as binary. The neuron's response to a given filtered stimulus value is then described in terms of a probability to observe a single spike rather than the number of spikes produced. To describe this function, we used the logistic function $P(\text{spike}|x) = [1 + \exp[-(x - \mu)/\nu]]^{-1}$ with two parameters, the threshold μ, which is the value of the filtered stimulus x

that yields 50% spiking probability, and the slope ν that determines when spiking probabilities deviate from either zero or one.

With binary responses, the slope describes noise present in the system (Fig. S1 *D* and *E*), because a shallow slope indicates that for certain x values spikes are elicited only on some fraction of trials. Note that this is true given the assumption that the cell is described by a linear filter followed by a static nonlinearity (LN model). Fractional spiking probabilities could also result from a cell with deterministic nonlinear dynamics such as a refractory period if it was fitted with an LN model that did not capture such nonlinear dynamics. However, within the LN model framework—a reasonable approximation to the firing rate of fast Off cells (14)—fractional spiking probabilities as found in the slope of the nonlinearity represent noise. The slope ν describes the effect of noise of variance $\nu^2/2$ added to the filtered stimulus x before it is passed through a deterministic thresholding function that equals 1 if $x > \mu$ and 0 if $x < \mu$. The deterministic thresholding function models the spike-generating nonlinearity, whereas the additive noise represents the cumulative effect of noise in the afferent circuitry; it may also include the component from the noise in the spike generation itself, which, however, is typically small compared with afferent noise (21).

Accounting for Differences in Numbers of Off/On Cell Types. We consider first the case where the two response functions have the same slope. Here, the optimal coding strategy changed as a function of the steepness of the slope, and two distinct behaviors emerged (Fig. 1*B*). In the regime with shallow slopes corresponding to large noise and ν, the maximally informative solution required that the two neurons have identical thresholds, with the threshold difference, $m = \mu_2 - \mu_1$, equal to zero. This regime corresponds to redundant encoding, where maximal information occurs through combining two noisy identical measurements on the signal, x. This indicates that even redundant encoding (22) can be an optimal strategy for populations of neurons. This solution stops being optimal when ν decreases below a certain critical value ν_c, where the optimal solution requires separate thresholds for the two neurons, and nonzero values of m become optimal. This highlights the importance of taking into account noise, something shown to be key for optimal encoding with transcription factors (23), as well as in cases where spikes are summed across time (24) or a neural population (25). Thus, the maximally informative solution for neural populations undergoes a sharp transition from one to two populations when the noise decreases below a critical value.

This transition between two different encoding schemes offers an explanation of a previously perplexing result. In salamanders all Off populations are heterogeneous, splitting into adapting and sensitizing populations. However, the On population is homogeneous, having a lower threshold (26) and displaying only adaptation (14). A possible explanation for the relative homogeneity of the On population could be that the noise in the On population is great enough that the optimal coding strategy would be redundant encoding. That predicts that the On response functions should have a shallower slope (larger ν) than the fast Off cells. We found this to be the case (Fig. 1 *C* and *D*). Not only do On cells have shallower slopes than Off cells as previously reported (27), but also their average slope lies above the critical point, placing them within the regime where redundant encoding is the optimal solution. Importantly, this match was obtained without any adjustable parameters, because the average firing rate extracted from experimental values uniquely determines the critical point. The error bars for the critical point value in Fig. 1*D* reflect the variation of the critical point values due to the different spiking rates recorded between pairs of neurons. The information maximization framework thus explains both the relative homogeneity of On cells, which do not split into separate classes, and the presence of two classes of Off neurons.

Optimal Threshold Differences Between Fast Off Populations. We then examined the information maximization framework by considering the case where the two response functions have unequal
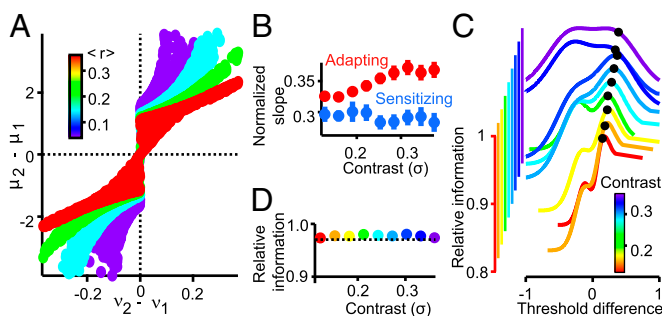
Fig. 2. Optimal dynamic range placement by fast Off populations. (*A*) Optimal threshold difference, $m = \mu_2 - \mu_1$, as a function of the differences in the slopes, $\nu_2 - \nu_1$, and the average rate $\langle r \rangle$ (color), which is constrained to a fixed value for each optimal solution. (*B*) Average slope values normalized by the contrast, σ, for the same set of adapting and sensitizing cells from Fig. 1*C*. (*C*) Information as a function of threshold difference between neurons. The values for slopes and a rate constraint were matched to measurements from a simultaneously recorded pair of adapting and sensitizing cells. The black solid circles show the measured threshold difference at each contrast. Each curve was normalized by the maximum information at that contrast (denoted by color). The curves are vertically offset from each for better visualization. (*D*) The average percentage of the maximum information reached for all cells pairs (*n* = 7) at each contrast is >97% (dotted line). Colors correspond to the colors from *C*. Error values, SEM, are obscured by the data points.

slopes. Fig. 2*A* shows that in optimal solutions positive differences between the two thresholds accompany positive differences between the corresponding slope values. That is, for the optimal solution the neuron with the higher threshold also has a response function with a shallower slope. This prediction was confirmed by experimental data (Fig. 2*B*), where across the full range of contrasts, sensitizing cells maintained steeper slopes than adapting cells. Because sensitizing cells also have lower thresholds (Fig. S1*C*), this coincides with the model prediction shown in Fig. 2*A*.

Up until this point we have used this theoretical framework to understand general features of the retinal data; however, it also has the capacity to determine the optimal threshold difference for individual pairs of fast Off ganglion cells. The corresponding analysis for an example pair of adapting and sensitizing cells is shown in Fig. 2*C*. Here, we fixed the slopes of the two response functions in the model to the measured values for this cell pair; we also fixed the average threshold across the two neurons to match the spike rate constraint determined from their measured response functions. The mutual information was then plotted as a function of the remaining parameter—the difference in thresholds between the two response functions. The measured threshold difference (black solid circles) was very close to the threshold difference that maximized the information about the input, for all contrasts. On average, fast Off cells had a threshold difference that provided >97% of the maximal amount of information across a large range of contrast distributions (Fig. 2*D*). Thus, these populations of neurons maintained the optimal position of their response functions even when the average spiking probability for the adapting and sensitizing cells increased with increasing contrast (Fig. S1).

Phase Transitions Within Maximally Informative Solutions. The behavior depicted in Fig. 1*B*, where the optimization function—here maximizing information—transforms from having a single optimal state to having two divergent optima, is one of the signatures of second-order phase transitions between different states of matter in physics (20). Systems near a phase transition exhibit a number of instabilities that are thought to increase the sensitivity of neural encoding to stimuli (17, 20, 28, 29). Characteristics of these instabilities depend upon the type of phase transition (20). To rigorously establish whether the transition that we find in our model corresponds to one of the known classes of phase transitions in physics we had to find the appropriate

correspondences between the key parameters that govern phase transitions in physics and the parameters used in our model. The theory of phase transitions draws its power from its ability to encompass diverse types of complex physical systems, mapping equivalent parameters from different physical systems onto each other. Even though some of the correspondences are not immediately intuitive, such as why density near the liquid–vapor critical point should correspond to magnetization in Ising magnetic systems, the theory provides a framework to explain why both of these quantities follow identical power law dependencies with respect to temperature (20).

As the first correspondence, we note that the observed states of matter come about through minimizing the free energy; in our solutions this corresponds to maximizing the information. Second, in physics, transitions occur with respect to temperature. In our case they occur as a function of the average slope, $\nu = (\nu_1 + \nu_2)/2$, of the two response functions, which describes the average noise in the neural responses. Third, in the Ising model, magnetization spontaneously appears below the critical temperature; here, the corresponding quantity is the threshold difference, $m = \mu_2 - \mu_1$, between the thresholds of the optimal response functions, which takes nonzero values below the critical noise level. Finally, to find the quantity analogous to the applied magnetic field, we note that an applied magnetic field induces nonzero magnetization even above the critical temperature. In our case, we find that a difference in the noise in the two response functions, $h = \nu_2 - \nu_1$, induces a nonzero optimal threshold difference between thresholds of the response functions over a broad range of noise levels (Fig. 2*A*). This suggests that the difference in the noise in the two response functions, h, or more generally the SD of ν values across the neural ensemble, is analogous to an applied magnetic field. We have also verified this correspondence quantitatively by showing that h, computed as the derivative of information with respect to magnetization, is linear with respect to $\nu_2 - \nu_1$ (Fig. S2). In Fig. 2*A*, the curves for different $\langle r \rangle$ values are plotted for the same average noise value ν. Because the critical noise value ν_c depends on the average firing rate $\langle r \rangle$ of the two neurons, the curves for different $\langle r \rangle$ are effectively placed at different distances from their respective critical points. This accounts for their spread.

Using these identified correspondences we can now determine the type of phase transition that occurs in our model with decreasing noise level. In most cases a phase transition is defined as a second-order transition when the singularities appear in the second derivatives of the optimization function. First, we examined the second derivative of information with respect to noise (Fig. S3*A*). This quantity is analogous to the specific heat, $C = \partial^2 I / \partial \nu^2$. We observe that C is largely independent from ν on each side of the transition with a sudden drop across the critical point. This is precisely the singular behavior expected based upon mean-field calculations for magnetization in magnetic systems, with smaller values observed above the critical temperature (20). Second, we found that the second derivative of the information with respect to h, $\chi = \partial^2 I / \partial h^2$, with h being the difference in slope of the response functions, displays a singularity at the critical point (Fig. S3*B*). This function χ is analogous to the magnetic susceptibility in magnetic transitions, which is the second derivative of the energy with respect to an applied magnetic field (20), and is sometimes interpreted as describing the system's sensitivity to external perturbations (29). Mean-field calculations indicate that this quantity should decay with an exponent of −1 as a function of temperature difference from its critical value (20). This matches the estimated exponent of −0.93 in our model (the difference from −1 reflects imprecision of numerical simulations). Thus, the transition we observe in neural circuits quantitatively matches behavior of the Ising model near its critical point.

At the critical point when $\nu = \nu_c$, the difference m between thresholds is zero and the difference in the noise of the two response functions h is also zero. The distance of the system from the critical point can thus be quantified by measuring how much the difference in thresholds m and the difference in the noise in the two response functions h differ from zero, as well as by the

deviation of average noise from its critical value $\nu - \nu_c$. Near the critical point, these three quantities are not independent, but rather scale as power-law functions of each other. Specifically, the mean-field Ising model calculation predicts that $m \propto |\nu - \nu_c|^\beta$ and $m \propto h^{1/\delta}$ with $\beta = 1/2$ and $\delta = 3$, respectively (20). We found that both of these relationships held true for our model of neural populations: with exponents $\beta = 0.47$ (Fig. S3C) and $1/\delta = 0.34$ (Fig. S3D), both of which closely matched their theoretical values of 0.5 and 0.33 (20), respectively. Therefore, by all metrics a system that maximizes information transmission in two populations falls into the class of models described by the Ising model of magnetism, which is called the Ising model universality class.

Fast Off Populations Remain Poised at the Critical Point. To examine where retinal circuits are positioned relative to the critical point, we measured where the threshold and slopes of each cell pair laid relative to the critical noise value ν_c. However, we needed to account for the fact that the average spike rate $\langle r \rangle$ differs for each pair of adapting/sensitizing neurons, and certain parameters such as ν_c vary with $\langle r \rangle$. Fortunately, the dependence of ν_c on $\langle r \rangle$ is stereotypic and smooth (Fig. S4A) and could, therefore, be normalized for each cell pair. In addition, in the equation that relates the threshold difference m to the critical noise value $m = A|\nu - \nu_c|^\beta$ (Fig. S3C) the dependence of the coefficient A on the average rate is also smooth (Fig. S4A). Finally, the scaling exponents β and $1/\delta$ do not depend on $\langle r \rangle$ (Fig. S4B). This makes it possible to transform the data into normalized coordinates where variables m, h, and ν_c do not depend on $\langle r \rangle$ (Fig. 3A). In these normalized coordinates, we can view all of the data from multiple pairs of cells relative to their respective critical points (Fig. 3B). We find that for all pairs of cells across the full range of contrasts the fast Off ganglion cells reside below the critical noise value, in the regime where it is optimal to split the encoding between response functions and two different thresholds. Thus, although the position of the critical point changes with mean spike rate, fast Off adapting and sensitizing cells maintain their response functions to stay below the critical point. We also note that whereas all of the results so far have been obtained for Gaussian signals, when more natural, non-Gaussian distributions are used as inputs, the parameters that characterize the singular behavior exhibited by the mutual information do not change (Fig. S5). Such robustness is expected for systems near phase transitions as part of their "universal" properties where divergent behavior near the critical point is not affected by changes in microscopic interactions. Similarly to free energy, the mutual information is computed by an average over the stimulus distribution. Changes in the higher-order parameters of the stimulus distribution alter the position of the critical point as a function of the spike rate but do not change its main characteristics: Differences in slope values still induce differences in thresholds and a single maximum in the information function splits into two for lower noise values.

Retina Scaling Exponents of Ising Model Universality Class. Remarkably, we find that even the deviations between the mean-field theory predictions and the experimental measurements in the retina matched the deviations observed in experiments on physical systems. Experimental measurements for the exponent β fall within a narrow range from 0.316 to 0.34 for all physical systems within the Ising universality class, including liquid–gas transitions in various substances, as well as ferromagnetic and antiferromagnetic transitions (20). Fitting our experimental data (Fig. 3B), we find a value, $\beta_{retina} = 0.39 \pm 0.12$ (Table 1). This value is consistent (taking into account the error bars) with the experimental observations in physical systems described by the 3D nearest-neighbor Ising model (20) and deviates from values expected for Ising models with nearest-neighbor interactions in other dimensions (Table S1). It is worth mentioning that, unlike in the physical systems where temperature is under complete experimental control, here we could not directly adjust ν, which limits our ability to precisely estimate the scaling exponents of our system from experimental data. For the exponent $1/\delta$, experiments in systems from the Ising model universality class produce values that are also shifted from the mean-field prediction of 1/3 to 0.204–0.217 (Table S1). Our data shift in the same direction, with a value $1/\delta_{retina} = 0.15 \pm 0.08$. This value is again consistent with the nearest-neighbor 3D Ising model (taking into account the error bars). Thus, the direction and magnitude of deviations that we observe here in the scaling exponents suggest that they have the same origins as the deviations observed in the physical systems from their mean-field values. This quantitative agreement in the way physical systems deviate from the mean-field predictions and the way the retina deviates from our mean-field theory-derived model further support the use of this simplifying framework in characterizing the critical and maximally informative behavior of the retina.

An important aspect of adaptive neural systems is that they adjust quickly to changing environments. In a natural sensory environment, changes in contrast occur often and unpredictably (30). To accommodate these changes in contrast, the slope of the response functions must adapt accordingly (2, 31, 32), necessitating an accompanying slow change in the threshold differences between the response functions to maintain optimal encoding. The fact that this process cannot and does not occur instantaneously (32) provides an additional factor that should influence response properties beyond a simple maximization of information transmission in one steady environment. In a stationary, or steady-state, regime, which corresponds to the analyses that we have carried out, neurons with sharper tuning functions (smaller ν) provide a greater amount of information (Fig. 1B). However, more narrow tuning functions, corresponding to a lower temperature, require more time to reach the optimal state. This phenomenon, known in physics as critical slowing down (20), occurs in the neural context because stimulus values that fall within the saturating region of the response function cannot be measured accurately
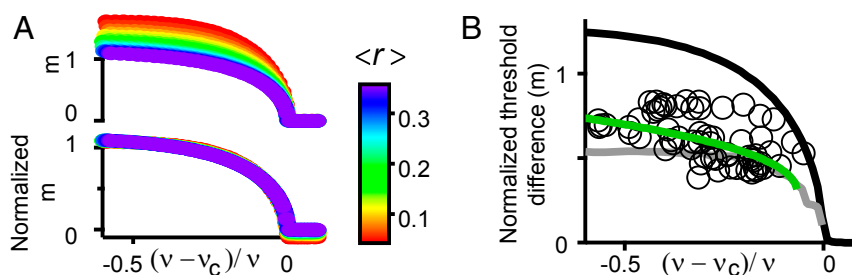

Fig. 3. Retinal populations adopt metastable states near the critical point. (A) In the model, the threshold difference m between cells is plotted against the normalized critical temperature, $(\nu - \nu_c)/\nu$ before (*Top*) and after (*Bottom*) normalizing for the average rate $\langle r \rangle$. (B) Threshold difference observed experimentally (black circles) normalized by the threshold difference that maximized the information (black line) and the spinodal line (gray curve), which delineates the regions between fast and slow dynamics near a critical point. Green line shows the average fit for the equation used to determine the scaling exponents (Table 1), using the average h value found in the data. These exponents match the Ising model universality class.

Table 1. Quantitative mapping between maximally informative solutions in the retina and Ising model of phase transitions in physics

Characterization	Ising magnetic systems	Maximally informative coding
Optimal states defined by	Minima of free energy	Maxima of information
Transitions occurs with respect to	Temperature	Input noise, average slopes of neural response functions
Symmetry broken below the critical temperature	Magnetization direction	Exchange symmetry between neurons*
Order parameter	Magnetization	SD of thresholds across a neural population†
Conjugate field	Applied magnetic field	SD of slopes from the mean across a neural population
Exponent with respect to temperature for $h = 0$	Mean-field value: 1/2	Mean-field value: 1/2
	Experiment: 0.316–0.327	Our experimental value: 0.39 ± 0.12
Critical isotherm exponent	Mean-field value: 1/3	Mean-field value: 1/3
	Experiment: 0.2–0.21	Our experimental value: 0.15 ± 0.08

The order parameter is the parameter that measures the degree to which solutions below the critical point deviate from the symmetric solution present above the critical point. The exponent β comes from the fit to the relationship between the threshold difference and the noise (Fig. S3C). The exponent $1/\delta$ comes from the fit to the relationship between the threshold difference and h, the difference between slopes of the two response functions (Fig. S3D). The values ($\pm$SD) for the scaling exponents (β and $1/\delta$) for our system were determined by a bootstrap fit to the seven pairs of adapting and sensitizing cells in the nine different contrasts, using the equation $A|\nu - \nu_c h|^\beta + Bh^{1/\delta}$ (Fig. 3B).
*See *SI Text*.
†See Fig. S2 for a quantitative verification of this definition.

and thus cannot trigger adaptation. Therefore, by using values of ν near ν_c, which provide less absolute information (Fig. 1B), the retina could be choosing an optimal strategy that it can reach quickly, rather than transmitting the largest possible amount of information at the expense of long adaptation times.

In support of these arguments, we found that although the observed values of threshold difference m achieved >97% of the maximum information (Fig. 2D), they were invariably smaller than the optimal separation (Fig. 3B). Tellingly, the data points lie near the so-called spinodal line that delineates the regions between fast and slow dynamics near a critical point (gray line in Fig. 3B). The dynamics necessary to increase m from the suboptimal point of zero (redundant encoding) to the spinodal line require only infinitesimal perturbations, whereas the dynamics of going from the spinodal line to the optimal line are expected to be slow, requiring large fluctuations (20). Thus, the steady-state value of the response curves was positioned to provide maximal information while remaining close to the boundary of a region that allowed for fast dynamics. Such properties of retinal responses are likely to facilitate adaptation and maximize information transmitted in nonstationary sensory environments.

Discussion

In this work we have shown that information maximization can explain the emergence and coordination of multiple cell types devoted to encoding of the same visual feature. The theory explains why neuronal types characterized by lower thresholds, such as sensitizing cells (Fig. S1C), have steeper tuning functions (Fig. 2B) compared with adapting cells that have higher thresholds. The maximally informative solutions undergo a bifurcation when the steepness of neural response function exceeds a certain value (Fig. 1B). The theory accounted for the bifurcation value without any adjustable parameters (Fig. 1 B–D). These bifurcations represent a robust phenomenon, because they persist even when the combinatorial representation of neural responses is simplified to a pooling rule where the identities of single neurons in the population are ignored (25) or when bell-shaped tuning curves are optimized to maximize Fisher information (33). We note that the separation of thresholds described by such bifurcations likely acts in addition to changes in the preferred stimulus features that are known to occur in response to changes in the effective noise level (34, 35).

The theory presented here does not explain directly why noise in the Off pathway should be less than that of the On pathway (26, 36). The functional advantages of such differences between the On and Off pathways have been ascribed to a greater prevalence of negative over positive contrasts in natural scenes (13). This excess of negative contrasts accounts for the larger number of Off neurons compared with On neurons (13, 37). Our results

add to these an explanation of why Off neurons split into two separate overlapping arrays of neurons with different thresholds as opposed to remaining a single denser array.

The description of why neurons in the brain form new classes from the information maximization perspective could be mapped to the theory of transitions between different phases of matter in physical systems (Table 1). Some of the connections make intuitive sense—noise in neural responses corresponds to temperature in physics—whereas other connections are more involved but respect the general properties of the information function, e.g., that it is an even function of m (Fig. 1B) and an odd function with respect to $h = \nu_2 - \nu_1$ (Fig. 2A). The set of correspondences described in Table 1 between the physical quantities and their counterparts in neural coding has yielded a clear quantitative match in the types of singularities that are observed in the two fields of science (Table 1 and Fig. S3 C and D). It could also be simply summarized by stating that in both physical and neural systems the optimization functions near their respective critical points behave approximately as

$$I \propto A(\nu - \nu_c)m^2 + Bm^4 + Chm, \quad [1]$$

where A, B, and C are constants. This expression corresponds to the Landau theory of phase transitions (20). The simple properties of this expansion formalize the argument that a neural circuit will be robust to changes that do not affect the control parameters, which consist of the mean ν and SD h of noise levels across the population. These control parameters in turn determine the optimal SD of thresholds m. As is the case for physical systems, Eq. 1 predicts a continuous second-order transition for $h = 0$ and a discontinuous first-order transition for $h \neq 0$. The signatures of the second-order transition are apparent in Fig. 1B and Fig. S3C where the threshold difference m varies continuously with ν for $h = 0$. The first-order transition reveals itself through the abrupt changes in m (Fig. 2A) when h changes sign for $\nu < \nu_c$. The dynamics of systems near first- and second-order transitions are affected not only by optimal solutions corresponding to information maximum but also by suboptimal "metastable" states. These metastable states likely reflect the fact that adaptation in neural systems does not occur instantaneously (32). The so-called spinodal line delineates the boundary between metastable and truly unstable states; it is also the curve along which the Fisher information about control parameters h and ν is maximized (38). The retinal data points fall near the spinodal curve, but on a side that corresponds to metastable as opposed to unstable solutions (Fig. 3B).

It should be emphasized that our theoretical derivations were obtained to highlight differences in the response functions of

NEUROSCIENCE

PHYSICS

neurons from different subpopulations, such as Off adapting and Off sensitizing cells, while ignoring differences in the response functions within each population. This corresponds to the mean-field approximation treatment of phase transitions in physics (20). It is well known that mean-field theory summarized by the expansion in Eq. **1** can capture qualitative features of system behavior near the critical point, but its predictions for scaling exponents deviate in a systematic way from experimental measurements. In physical systems, these discrepancies have been resolved through the development of the renormalization group (RG) theory that builds upon on mean-field approximation but then takes into account the fluctuations in control parameters across the array (20). In this case, a future goal will be to use RG theory to take into account small differences in response parameters across the retinal array and the fact that inputs to individual cells are not perfectly correlated and differ among each other.

Our measurements of scaling exponents in the retina matched experimental measurements in physical systems that correspond to the 3D Ising model (Fig. 3*B* and Table 1), ruling out matches to nearest-neighbor Ising models of other dimensions (Table S1). Also consistent with a 3D Ising model is the previous observation (15) of a cusp-like singularity in the quantity analogous to the specific heat. Whereas the mean-field theory predicts that specific heat changes discontinuously across the transition (Fig. S3*A*), RG corrections predict a cusp in agreement with experiments in physical systems. One may wonder how the match to the 3D exponent can be consistent with retinal ganglion cells arranged in a 2D array. It turns out that, in terms of critical exponents, an Ising model with interactions that extend beyond nearest neighbors is more equivalent to a higher-dimensional model than one based on nearest-neighbor interactions, with the dimensionality depending on the spatial decay of the interaction strength (39). Given that retinal receptive fields are centered on a 2D lattice, our finding that critical exponents match the 3D Ising model implies that the effective interaction strength, which can be estimated using models from ref. 40, decreases with

distance r between receptive field center positions as $r^{-3.309}$. This parameter-free prediction can be tested in future studies.

The mapping that makes it possible to apply insights from the theory of phase transitions to neural systems relies on a slight but important change in perspective on what are analogous elements between neural and magnetic systems. Previous studies have typically identified single spins and spikes in single neurons as analogous, because both are binary. We find that spins are more analogous to differences in responses across neurons, which in turns highlights the relative nature of the neural code. Overall, it is fitting that a theory developed in physics to tackle the case of complex systems with many interactive degrees of freedom can also offer insights into the function of neural circuits. Perhaps broader application of the ideas described here could explain the existence of other classes of cells throughout the brain.

Materials and Methods

Recordings were done from larval tiger salamander retinas, using a 60-electrode array as previously described (14). Experiments were performed in accordance with Stanford University Animal Care and Use guidelines. Stimuli were uniform field with a constant mean intensity of 10 mW/m² and were drawn from a Gaussian distribution and presented at 30 Hz. Contrast was defined as the ratio between the SD and the mean of the intensity distribution. Neurons were probed with nine uniformly spaced contrasts from 0.12 to 0.36 (*SI Methods*).

1. Laughlin S (1981) A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch C* 36(9-10):910–912.
2. Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. *Neuron* 26(3):695–702.
3. Pitkow X, Meister M (2012) Decorrelation and efficient coding by retinal ganglion cells. *Nat Neurosci* 15(4):628–635.
4. Atick JJ, Redlich AN (1990) Towards a theory of early visual processing. *Neural Comput* 2:308–320.
5. Haft M, van Hemmen JL (1998) Theory and implementation of infomax filters for the retina. *Network* 9(1):39–71.
6. Doi E, et al. (2012) Efficient coding of spatial information in the primate retina. *J Neurosci* 32(46):16256–16264.
7. Liu YS, Stevens CF, Sharpee TO (2009) Predictable irregularities in retinal receptive fields. *Proc Natl Acad Sci USA* 106(38):16499–16504.
8. Borghuis BG, Ratliff CP, Smith RG, Sterling P, Balasubramanian V (2008) Design of a neuronal array. *J Neurosci* 28(12):3178–3189.
9. Zhaoping L (2006) Theoretical understanding of the early visual processes by data compression and data selection. *Network* 17(4):301–334.
10. Garrigan P, et al. (2010) Design of a trichromatic cone array. *PLoS Comput Biol* 6(2):e1000677.
11. Gjorgjieva J, Sompolinsky H, Meister M (2014) Benefits of pathway splitting in sensory coding. *J Neurosci* 34(36):12127–12144.
12. Balasubramanian V, Sterling P (2009) Receptive fields and functional architecture in the retina. *J Physiol* 587(Pt 12):2753–2767.
13. Ratliff CP, Borghuis BG, Kao YH, Sterling P, Balasubramanian V (2010) Retina is structured to process an excess of darkness in natural scenes. *Proc Natl Acad Sci USA* 107(40):17368–17373.
14. Kastner DB, Baccus SA (2011) Coordinated dynamic encoding in the retina using opposing forms of plasticity. *Nat Neurosci* 14(10):1317–1322.
15. Tkacik G, Schneidman E, Berry MJ, Bialek W (2009) Spin glass models for a network of real neurons. arXiv:0912.5409.
16. Mora T, Bialek W (2011) Are biological systems poised at criticality? *J Stat Phys* 144:268–302.
17. Shew WL, Yang H, Yu S, Roy R, Plenz D (2011) Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J Neurosci* 31(1):55–63.
18. Friedman N, et al. (2012) Universal critical dynamics in high resolution neuronal avalanche data. *Phys Rev Lett* 108(20):208102.
19. Beggs JM, Timme N (2012) Being critical of criticality in the brain. *Front Physiol* 3:163.
20. Stanley HE (1971) *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ Press, New York).
21. Ala-Laurila P, Greschner M, Chichilnisky EJ, Rieke F (2011) Cone photoreceptor contributions to noise and correlations in the retinal output. *Nat Neurosci* 14(10):1309–1316.
22. Puchalla JL, Schneidman E, Harris RA, Berry MJ (2005) Redundancy in the population code of the retina. *Neuron* 46(3):493–504.
23. Tkacik G, Callan CG, Jr, Bialek W (2008) Information flow and optimization in transcriptional regulation. *Proc Natl Acad Sci USA* 105(34):12265–12270.
24. Nikitin AP, Stocks NG, Morse RP, McDonnell MD (2009) Neural population coding is optimized by discrete tuning curves. *Phys Rev Lett* 103(13):138101.
25. McDonnell MD, Stocks NG, Pearce CEM, Abbott D (2006) Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance. *Phys Lett A* 352(3):183–189.
26. Kastner DB, Baccus SA (2013) Spatial segregation of adaptation and predictive sensitization in retinal ganglion cells. *Neuron* 79(3):541–554.
27. Zaghloul KA, Boahen K, Demb JB (2003) Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *J Neurosci* 23(7):2645–2654.
28. Kinouchi O, Copelli M (2006) Optimal dynamic range of excitable networks at criticality. *Nat Phys* 2:348–351.
29. Yu S, Yang H, Shriki O, Plenz D (2013) Universal organization of resting brain activity at the thermodynamic critical point. *Front Syst Neurosci* 7:42.
30. Frazor RA, Geisler WS (2006) Local luminance and contrast in natural images. *Vision Res* 46(10):1585–1598.
31. Baccus SA, Meister M (2002) Fast and slow contrast adaptation in retinal circuitry. *Neuron* 36(5):909–919.
32. Ozuysal Y, Baccus SA (2012) Linking the computational structure of variance adaptation to biophysical mechanisms. *Neuron* 73(5):1002–1015.
33. Harper NS, McAlpine D (2004) Optimal neural population coding of an auditory spatial cue. *Nature* 430(7000):682–686.
34. Karklin Y, Simoncelli EP (2011) Efficient coding of natural images with a population of noisy Linear-Nonlinear neurons. *NIPS* 24:999–1007.
35. Carlson NL, Ming VL, Deweese MR (2012) Sparse codes for speech predict spectrotemporal receptive fields in the inferior colliculus. *PLoS Comput Biol* 8(7):e1002594.
36. Chichilnisky EJ, Kalmar RS (2002) Functional asymmetries in ON and OFF ganglion cells of primate retina. *J Neurosci* 22(7):2737–2747.
37. Segev R, Puchalla J, Berry MJ, 2nd (2006) Functional organization of ganglion cells in the salamander retina. *J Neurophysiol* 95(4):2277–2292.
38. Brody D, Rivier N (1995) Geometrical aspects of statistical mechanics. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 51(2):1006–1011.
39. Angelini MC, Parisi G, Ricci-Tersenghi F (2014) Relations between short-range and long-range Ising models. *Phys Rev E Stat Nonlin Soft Matter Phys* 89(6):062120.
40. Schneidman E, Berry MJ, 2nd, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440(7087):1007–1012.