Weak pairwise correlations imply strongly correlated network states in a neural population

Elad Schneidman,[a−c] Michael J. Berry II,[b] Ronen Segev[b] and William Bialek[a,c]

[a]*Joseph Henry Laboratories of Physics,* [b]*Department of Molecular Biology,* and [c]*Lewis–Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544 USA*

(Dated: December 6, 2005)

Biological networks have so many possible states that exhaustive sampling is impossible. Successful analysis thus depends on simplifying hypotheses, but experiments on many systems hint that complicated, higher order interactions among large groups of elements play an important role. In the vertebrate retina, we show that weak correlations between pairs of neurons coexist with strongly collective behavior in the responses of ten or more neurons. Surprisingly, we find that this collective behavior is described quantitatively by models that capture the observed pairwise correlations but assume no higher order interactions. These maximum entropy models are equivalent to Ising models, and predict that larger networks are completely dominated by correlation effects. This suggests that the neural code has associative or error- correcting properties, and we provide preliminary evidence for such behavior. As a first test for the generality of these ideas, we show that similar results are obtained from networks of cultured cortical neurons.

Much of what we know about biological networks has been learned by studying one element at a time— recording the electrical activity of single neurons, the expression levels of single genes or the concentrations of individual metabolites. On the other hand, important aspects of biological function must be shared among many elements (1; 2; 3; 4). As a first step beyond the analysis of elements in isolation, much attention has been focused on the pairwise correlation properties of these elements, both in networks of neurons (5; 6; 7; 8; 9; 10; 11; 12; 13) and in networks of genes (14; 15; 16). But given a characterization of pairwise correlations, what can we really say about the whole network? How can we tell if inferences from a pairwise analysis are correct, or if they are defeated by higher order interactions among triplets, quadruplets, ... of elements? If these effects are important, how can we escape from the 'curse of dimensionality' that arises because there are exponentially many possibilities for such terms?

Here we address these questions in the context of the vertebrate retina, where it is possible to make long, stable recordings from many neurons simultaneously as the system responds to complex, naturalistic inputs (17; 18; 19; 20). We compare the correlation properties of cell pairs with the collective behavior in larger groups of cells, and find that the minimal model which incorporates the pairwise correlations provides strikingly accurate but non–trivial predictions of the collective effects. These minimal models are equivalent to the Ising model in statistical physics, and this mapping allows us to explore the properties of larger networks, in particular their capacity for error–correcting representations of incoming sensory data.[1]

I. THE SCALE OF CORRELATIONS

Throughout the nervous system, individual elements communicate by generating discrete pulses termed action potentials or spikes (21). If we look through a window of fixed time resolution $\Delta\tau$, then for small $\Delta\tau$ these responses are binary—either the cell spikes ('1') or it doesn't ('0'). Although some pairs of cells have very strong correlations, most correlations are weak, so that the probability of seeing synchronous spikes is almost equal to the product of the probabilities of seeing the individual spikes; nonetheless, these weak correlations are statistically significant for most if not all pairs of nearby cells. All of these features are illustrated quantitatively by an experiment on the salamander retina (Fig. 1), where we record simultaneously from 40 retinal ganglion cells as they respond to movies taken from a natural setting (see Methods). Mean spike rates range from 0.3 to 4.5 spikes/s and the correlations between cells have structure on the scale of $\Delta\tau = 20\,\text{ms}$. Using this window to define binary variables, we find correlation coefficients that range from -0.03 to 0.5, with ninety percent of the results falling in the narrow range $C \in [-0.02, +0.1]$.

The small values of the correlation coefficients suggest an approximation in which the cells are completely independent. For most pairs, this is true with a precision of a few percent, but if we extrapolate this approximation to the whole population of 40 cells, it fails disastrously. In Fig. 1e we show the probability $P(K)$ that K of these cells generate a spike in the same small window of duration $\Delta\tau$. If the cells were independent, $P(K)$ would approximate the Poisson distribution, while the true dis-

---

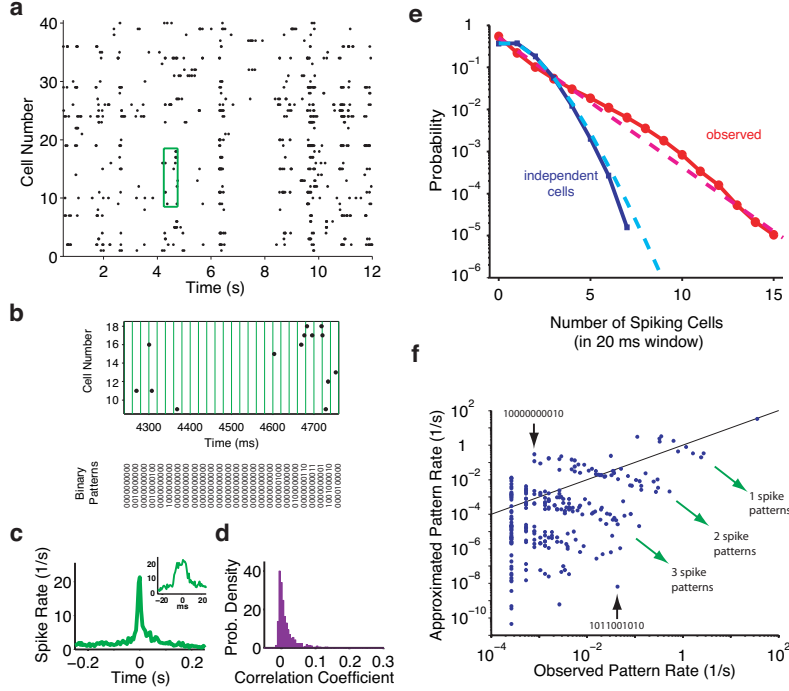[1] A preliminary account of this work was presented at

FIG. 1 Weak pairwise cross–correlations and the failure of the independent approximation. (a) A segment of the simultaneous responses of 40 retinal ganglion cells in the salamander to a natural movie clip. Each dot represents the time of an action potential. (b) Discretization of population spike trains into a binary pattern is shown for the green boxed area in panel a: spike trains are binned using $\Delta\tau = 20\,\mathrm{ms}$ windows (top panel), into binary sequences of spiking (1) and non–spiking (0); in the rare cases where there is more than one spike in a bin, we denote it as '1'. Every string (bottom panel) describes the activity pattern of the cells on a give time point. For clarity, 10 out of 40 cells are shown. (c) Example cross–correlogram between two neurons with strong correlations; the average firing rate of one cell is plotted relative to time at which the other cell spikes. Inset shows the same cross–correlogram on an expanded time scale. (d) Histogram of correlation coefficients for all pairs of 40 cells from panel a. We discretize the neural response into binary (spike/no spike) variables for each cell, using $\Delta\tau = 20\,\mathrm{ms}$ bins as in (b), and then compute the correlation coefficients among these variables. Because the data sets we consider here are very large ($\sim 1\,\mathrm{hour}$), the threshold for statistical significance of the individual correlation coefficients is well below $|C| = 0.01$. (e) Probability distribution of synchronous spiking events in the 40 cell population in response to a long natural movie (red) approximates an exponential (dashed red). The distribution of synchronous events for the same 40 cells after shuffling each cell's spike train to eliminate all correlations (blue), compared to the Poisson distribution (dashed light blue). (f) The rate of occurrence of each pattern predicted if all cells are independent is plotted against the measured rate. Each dot stands for one of the $2^{10} = 1024$ possible binary activity patterns for 10 cells. Black line shows equality. Two examples of extreme mis–estimation of the actual pattern rate by the independent model are highlighted (see text). The logarithmic scale emphasizes the large dynamic range and large errors of the independent approximation; less clear is that the (very accurately measured) probability of the response 0000000000 is underestimated by $\sim 7\%$.

tribution is nearly exponential. The probability of $K = 7$ cells spiking together is $\sim 10^3\times$ larger than expected in the independent model, and at $K = 10$ the discrepancy is $\sim 10^5$: 10–spike events occur several times per minute in this small patch of the retina, while the independent model predicts that they should occur once every three weeks.

The discrepancy between the independent model and the actual data is even more clear if we look at particular patterns of response across the population. Choosing $N = 10$ cells out the 40, we can form an N–letter binary word to describe the instantaneous state of the network, as in Fig 1b. The independent model makes simple pre-

dictions for the rate at which each such word should occur, and Fig.1f shows these predictions as a scatter plot against the actual rate at which the words occur in the experiment. At one extreme, the word 1011001010 occurs once per minute, while the independent model predicts that this should occur once per three years. Conversely, the word 1000000010 is predicted to occur once per three seconds, while in fact it occurred only three times in the course of an hour. Although these are extreme cases, the independent model makes order–of–magnitude errors even for the rates of very common patterns of activity, such as a single cell generating a spike while all others are silent. Indeed, in the scatter plot of predicted vs.

observed rates we see clusters corresponding to different total numbers of spikes, but within each cluster the predictions and observations are strongly anti–correlated.

We conclude that weak correlations among pairs of neurons coexist with strong correlations in the states of the population as a whole. One possible explanation is that there are specific multi–neuron correlations, whether driven by the stimulus or intrinsic to the network, which simply are not measured by looking at pairs of cells. Searching for such higher order effects presents many challenges (22; 23; 24). Another scenario is that small correlations among very many pairs could add up to a strong effect on the network as a whole. If correct, this would be an enormous simplification in our description of the network dynamics.

II. MINIMAL CONSEQUENCES OF PAIRWISE CORRELATIONS

To describe the network as a whole we need to write down a probability distribution for the 2^N binary words corresponding to patterns of spiking and silence in the population. The pairwise correlations tell us something about this distribution, but there are an infinite number of models that are consistent with a given set of pairwise correlations. The difficulty thus is to find a distribution that is consistent *only* with the measured correlations, and does not implicitly assume the existence of unmeasured higher–order interactions. Since the entropy of a distribution measures the randomness or lack of interaction among different variables (25), this minimally structured distribution that we are looking for is the maximum entropy distribution (26) consistent with the measured properties of individual cells and cell pairs (27).

We recall that maximum entropy models have a close connection to statistical mechanics: physical systems in thermal equilibrium are described by the Boltzmann distribution, which has the maximum possible entropy given the mean energy of the system (26; 28). Thus any maximum entropy probability distribution defines an energy function for the system we are studying, and we will see that the energy function relevant for our problem is an Ising model. Ising models have been discussed extensively as models for neural networks (29; 30), but in these discussions the model arose from specific hypotheses about the network dynamics. Here the Ising model is forced upon us as the least structured model that is consistent with measured spike rates and pairwise correlations; we emphasize that this is not an analogy or a metaphor, but rather an exact mapping.

Whether we view the maximum entropy model through its analogy with statistical physics or simply as a model to be constructed numerically from the data (see Methods), we need meaningful ways of assessing whether this model is correct. Generally, for a network of N elements, we can define maximum entropy distributions P_K that are consistent with all K^{th}–order correlations for any

$K = 1, 2, \cdots, N$ (27). These distributions form a hierarchy, from $K = 1$ where all elements are independent, up to $K = N$, which is an exact description that allows arbitrarily complex interactions; their entropies S_K decrease monotonically toward the true entropy S: $S_1 \geq S_2 \geq \cdots \geq S_N = S$. The entropy difference or multi–information $I_N = S_1 - S_N$ measures the total amount of correlation in the network, independent of whether it arises from pairwise, triplet, or more complex correlations (31). The contribution of the Kth–order correlation is $I_K = S_{K-1} - S_K$ and is always positive (more correlation always reduces the entropy); I_N is the sum of all the I_K (27). Therefore, the question of whether pairwise correlations provide an effective description of the system becomes the question of whether the reduction in entropy that comes from these correlations, $I_2 = S_1 - S_2$, captures all or most of the multi–information I_N.

III. ARE PAIRWISE CORRELATIONS ENOUGH?

Figure 2 shows the predictions of the maximum entropy model P_2 consistent with pairwise correlations in populations of $N = 10$ cells. Looking in detail at the patterns of spiking and silence of one group of 10 cells, we see that the predicted rates for different binary words are tightly correlated with the observed rates over a very large dynamic range, so that the dramatic failures of the independent model have been overcome (Fig 2a).

With 40 cells we can choose many different populations of 10 cells, and in each case we find that the predicted and observed distributions of words are very similar. This is quantified by the Jensen–Shannon divergence D_{JS} (34), which essentially measures the inverse of the number of independent samples we would need in order to be sure that the two distributions were different. While the true distribution of responses is not exactly the same as that predicted by the maximum entropy model, it typically would take thousands of independent samples to distinguish reliably between them, two orders magnitude more than in the independent model (Fig. 2b).

The success of the pairwise maximum entropy models in capturing the correlation structure of the network is summarized by the fraction $I_2/I_N \sim 90\%$ (Fig 2c). This ratio is larger when I_N itself is larger, so that the pairwise model is more effective in describing populations of cells with stronger correlations, and the ability of this model to capture $\sim 90\%$ of the multi–information holds independent of many details (Fig 2d): We can vary the particular natural movies shown to the retina, use an artificial movie, change the size of the bins $\Delta\tau$ used to define the binary responses, the number of neurons N that we analyze, and even shift from a lower vertebrate (salamander) to a mammalian (guinea pig) retina. Finally, the correlation structure in a network of cultured cortical neurons (41) can be captured by the pairwise model with similar accuracy.
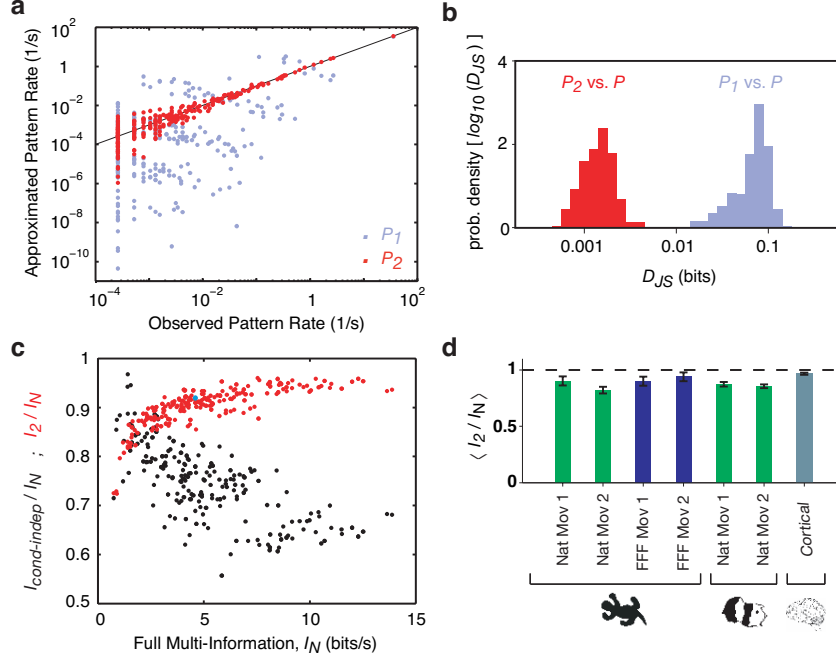
The maximum entropy model describes the correla-

FIG. 2 A maximum entropy model including all pairwise interactions gives an excellent approximation of the full network correlation structure. (a) Using the same group of 10 cells from Fig. 1, the rate of occurrence of each firing pattern predicted from the the maximum entropy model P_2 that takes into account all pairwise correlations is plotted against the measured rate (red dots). The rates of commonly occurring patterns are predicted with better than 10% accuracy, and scatter between predictions and observations is confined largely to rare events for which the measurement of rates is itself uncertain. For comparison, the independent model P_1 is also plotted (from Fig. 1f; gray points). Black line shows equality. (b) Histogram of Jensen–Shannon divergences between the actual probability distribution of activity patterns in 10 cell groups and the models P_1 (gray) and P_2 (red); data from 250 groups. Note that the histograms are plotted on a logarithmic scale, because the two cases have values differing by two orders of magnitude. (c) Fraction of full network correlation in 10 cell groups that is captured by the maximum entropy model of 2nd order, I_2/I_N, plotted as a function of the full network correlation, measured by the multi–information I_N (red dots). The multi–information values are multiplied by $1/\Delta\tau$ to give bin–independent units. Every dot stands for one group of 10 cells. The 10 cell group featured in panel a is shown as an light blue dot. For the same sets of 10 cells, the fraction of information of full network correlation in 10 cell groups that is captured by the conditional independence model, $I_{\mathrm{cond-indep}}/I_N$, is shown in black (see text). (d) Average values of I_2/I_N from 250 different 10 cell groups. Results are shown for different movies, for different species, and for cultured cortical networks; error bars show standard errors of the mean. Similar results are obtained on changing N and $\Delta\tau$.

tion structure of the network activity without assumptions about its mechanistic origin. A more traditional approach has been to dissect the correlations into contributions that are intrinsic to the network and those that are driven by the visual stimulus. The simplest model in this view is one in which cells spike independently in response to their input, so that all correlations are generated by covariations of the individual cells' firing rates (35). Although there may be situations in which conditional independence is a good approximation, Fig 2c shows that this model is less effective than the maximum entropy model in capturing the multi–information for 232/250 groups of 10 neurons. The hypothesis of conditional independence is consistently less effective in capturing the structure of more strongly correlated groups of cells, which is opposite to the behavior of the maximum entropy model. Note also that because each cell

has its own spike rate, potentially different at each moment in time, the conditionally independent model has has $NT/\Delta\tau$ parameters, where T is the duration of the stimulus movie; in our case $NT/\Delta\tau \sim 10^4$, in contrast to the $N(N+1)/2 = 55$ parameters of the maximum entropy model. Finally, while the maximum entropy model can be constructed solely from the observed correlations among neurons, the conditionally independent model requires explicit access to repeated presentations of the visual stimulus. Thus while the central nervous system could learn the maximum entropy model from the data provided by the retina alone, the conditionally independent model is not biologically realistic in this sense.

We conclude that although the pairwise correlations are small and the multi–neuron deviations from independence are large, the maximum entropy model consistent with the pairwise correlations captures almost all of the

structure in the distribution of responses from the full population of neurons. Thus, the weak pairwise correlations *imply* strongly correlated states. To understand how this happens, it is useful to look at the mathematical structure of the maximum entropy distribution.

IV. ISING MODELS, REVISITED

We recall that the maximum entropy distribution consistent with a known average energy $\langle E \rangle$ is the Boltzmann distribution, $P \propto \exp(-E/k_B T)$. This generalizes so that if we know the average values of many variables f_μ describing the system, then the maximum entropy distribution is $P \propto \exp(-\sum_\mu \lambda_\mu f_\mu)$, where there is a separate Lagrange multiplier λ_μ for each constraint (26; 28). In our case, we are given the average probability of a spike in each cell and the correlations among all pairs. If we represent the activity of cell i by a variable $\sigma_i = \pm 1$, where $+1$ stands for spiking and -1 stands for silence, then these constraints are equivalent to fixing the average of each σ_i and the averages of all products $\sigma_i \sigma_j$, respectively. The resulting maximum entropy distribution is

$$
\begin{aligned}
&P_2(\sigma_1, \sigma_2, \cdots, \sigma_N) \\
&= \frac{1}{Z} \exp\left[\sum_i h_i \sigma_i + \frac{1}{2} \sum_{i \neq j} J_{ij} \sigma_i \sigma_j \right],
\end{aligned} \tag{1}
$$

where the Lagrange multipliers $\{h_i, J_{ij}\}$ have to be chosen so that the averages $\{\langle \sigma_i \rangle, \langle \sigma_i \sigma_j \rangle\}$ in this distribution agree with experiment. This is the Ising model (28), where the σ_i are spins, the h_i are local magnetic fields acting on each spin, and the J_{ij} are the exchange interactions; note that $h > 0$ favors spiking and $J > 0$ favors positive correlations.

Figure 3 shows the parameters $\{h_i, J_{ij}\}$ for a particular group of ten cells, as well as the distributions of parameters for many such groups. Most cells have a negative local field, which biases them toward silence. Figures 3c and d illustrates the non–trivial relationship between the pairwise interaction strengths J_{ij} and the observed pairwise correlations. Correlations can be induced between two cells in the absence of direct interaction when both cells interact with a common neighbor (Fig 3c, left). Alternatively, if the interactions among pairs of cells have alternating signs, then it is possible to have "frustration" among triplets of cells destroy the correlations between two cells even when they have a strong positive interaction (Fig 3c, right). Roughly 40% of all triangles are frustrated in this way, and we return to this below.

We can rewrite Eq (1) exactly by saying that each neuron or spin σ_i experiences an effective magnetic field that includes the local field or intrinsic bias h_i and a contribution from interactions with all the other spins, $h_i^{\text{int}} = \frac{1}{2} \sum_{j \neq i} J_{ij} \sigma_j$; note that h_i^{int} depends on whether the other cells are spiking or silent. The intrinsic bias dominates in small groups of cells, but as we look to
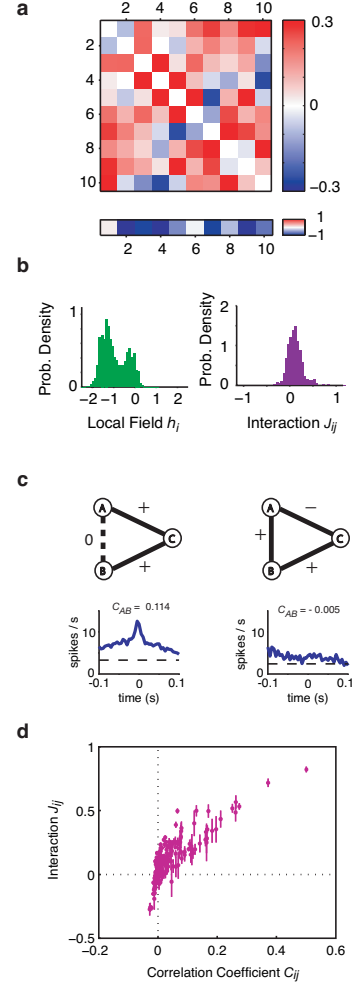

FIG. 3 Pairwise interactions and individual cell biases, as in Eq. (1). (a) Example of the pairwise interactions J_{ij} (above) and bias values (or local fields) h_i (below) for one group of 10 cells. (b) Histograms of h_i and J_{ij} values from 250 different groups of 10 cells. (c) Two examples of 3 cells within a group of 10. At left, cells A and B have almost no interaction ($J_{AB} = -0.02$), but cell C is very strongly interacting with both A and B ($J_{AC} = 0.52, J_{BC} = 0.70$), so that cells A and B exhibit strong correlation, as shown by their cross–correlogram (bottom panel; $C_{AB} = 0.11$). At right, a "frustrated" triplet, in which cells A and B have a significant positive interaction ($J_{AB} = 0.13$), as do cells B and C ($J_{BC} = 0.09$), but A and C have a significant negative interaction ($J_{AC} = -0.11$). As a result, there is no clear correlation between cells A and B, as shown by their cross–correlogram (bottom panel; $C_{AB} = -0.005$). (d) Interaction strength J_{ij} plotted against the correlation coefficient C_{ij}; each point shows the value for one cell pair averaged over many different groups of neighboring cells (190 pairs from 250 groups), and error bars show standard deviations.
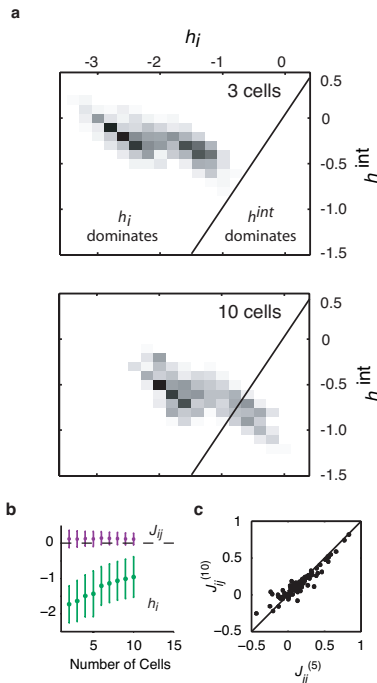
FIG. 4 Interactions and local fields in networks of different size. (a) Grayscale density map of the distribution of effective interaction fields experienced by a single cell h_i^{int} versus its own bias or local field h_i (see text); distribution formed over network configurations at each point in time during a natural movie for n=1140 3–cell groups (top panel) and n=250 10–cell groups (bottom panel). Black line shows the boundary between dominance of local fields vs. interactions. (b) Mean interactions J_{ij} and local fields h_i describing groups of N cells, with error bars showing standard deviations across multiple groups. While local fields decline significantly with increasing network size, interaction strengths do not. (c) Pairwise interaction in a network of 10 cells J_{ij}^{10} plotted against the interaction values of the same pair in a sub-network containing only 5 cells J_{ij}^5. Line shows equality.

larger networks, the fact that almost all of the $\sim N^2$ pairs of cells are significantly (if weakly) interacting shifts the balance so that the typical values of the intrinsic bias are reduced while the effective field contributed by the other cells has increased (Fig. 4). Intuitively, although the influence of one cell on another is small, the summed effect of many cells on one cell can be large, so that states of the whole network are far from those predicted by the independent approximation.

V. LARGER NETWORKS AND ERROR CORRECTION

Groups of $N = 10$ cells are large enough to reveal dramatic departures from independence, but small enough that we can directly sample the relevant probability distributions. What happens at larger N? In general we

expect that the total capacity of the network to represent its sensory inputs should grow in proportion to the number of neurons N. This is the usual thermodynamic limit in statistical mechanics, where energy and entropy are proportional to system size (28). But this behavior is not guaranteed when all elements of the system interact with each other. In the Ising model, it is known that if all pairs of spins (here, cells) interact significantly, then to recover the thermodynamic limit the typical size of the interactions J_{ij} must decrease with N (30; 36). Although we have not analyzed very large networks, we see no signs of significant changes in J with growing N (Fig 4b, c).

In a physical system, the maximum entropy distribution is equivalent to the Boltzmann distribution $P \propto \exp(-E/k_B T)$, and the behavior of the system depends on the temperature T. For the network of neurons, there is no real temperature, but the statistical mechanics of the Ising model predicts that when all pairs of elements interact, increasing the number of elements while fixing the typical strength of interactions is equivalent to lowering the temperature T in a physical system of fixed size N. This mapping predicts that correlations will be even more important in larger groups of neurons.

We can see signs of strong correlation emerging by looking at the entropy and multi–information in networks of different sizes. If all cells were independent, the entropy would be S_1, exactly proportional to N. For weak correlations, we can solve the Ising model in perturbation theory to show that the multi–information I_N is the sum of mutual information terms between all pairs of cells and hence $I_N \propto N^2$. This is in agreement with the empirically estimated I_N up to $N = 15$, the largest value for which direct sampling of the data provides a good estimate (Fig 5a), and Monte Carlo simulations of the maximum entropy models suggest that this agreement extends up to the full population of $N = 40$ neurons in our experiment (G Tkačik, ES, RS, MJB & WB, unpublished). Were this pattern to continue, at $N \sim 200$ cells I_N would become equal to the independent entropy S_1, and the true entropy $S_N = S_1 - I_N$ would vanish as the system "froze." Because we see variable firing patterns of all the cells, we know that literal freezing of the network into a single state doesn't happen. On the other hand, networks of $N \sim 200$ cells must be very strongly ordered. Interestingly, experiments indicate that a correlated patch on the retina has roughly this size: the strongest correlations are found for cells within $\sim 200\,\mu$m of each other, and this area contains ~ 175 ganglion cells in the salamander (19).

Because the interactions J_{ij} have different signs, frustration can prevent the freezing of the system into a single state. Instead there will be multiple states that are local minima of the effective energy function, as in spin glasses (36). If the number of minimum energy patterns is not too small, then the system retains a significant representational capacity. If the number of patterns is not too large, then observing only some of the cells in the network
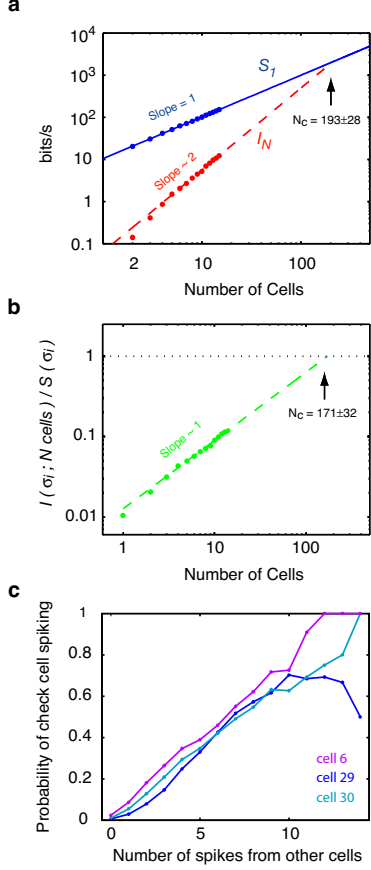
FIG. 5 Extrapolation to larger networks. (a) Average independent cell entropy S_1 and network multi-information I_N, multipled by $1/\Delta\tau$ to give bin–independent rates, vs. number of cells in the network N. Theoretically we expect $I_N \propto N(N-1)$ for small N; the best fit is $I_N \propto N^{1.98\pm0.04}$. Extrapolating (dashed line) defines a critical network size N_c, where I_N would be equal to S_1. (b) Information that N cells provide about the activity of the $N+1^{\mathrm{st}}$ cell, plotted as a fraction of that cell's entropy, vs. network size N; each point is the average value for many different groups of cells. Extrapolation to larger networks (dashed line, slope $= 1.017\pm0.052$) defines another critical network size N_c, where one would get perfect error–correction or prediction of the state of a single cell from the activity of the rest of the network. (c) Examples of 'check cells,' for which the probability of spiking is an almost perfectly linear encoding of the number of spikes generated by the other cells in the network. Cell numbers as in Fig 1.

is sufficient to identify the whole pattern uniquely, just as in the Hopfield model of associative memory (29). Thus the system would have a holographic or error–correcting property, so that an observer who has access only to a fraction of the neurons would nonetheless be able to reconstruct the activity of the whole population.

We can see suggestions of this error–correcting property by asking directly how much the knowledge of activity in N cells tells us about the whether the $N+1^{\mathrm{st}}$ cell will spike (Fig 5b). Our uncertainty about the state of one cell is reduced in proportion to the number of cells that we examine, and if this trend continues then again at $N \sim 200$ all uncertainty would vanish. Alternatively, we can look for particular kinds of error correction. In our population of 40 cells we have found three cells for which the probability of spiking is an almost perfectly linear encoding of the number of spikes generated by the other cells in the network (Fig 5c); by observing the activity of these "check cells" we can estimate how many spikes are generated by the network as a whole even before we observe any of the other cells' responses.

VI. DISCUSSION

Returning to the general questions raised at the outset, we have seen that the maximum entropy principle provides a unique candidate model for the whole network that is consistent with observations on pairs of elements but makes no additional assumptions. Despite the opportunity for higher–order interactions in the retina, this model captures more than 90% of the structure in the detailed patterns of spikes and silence in the network, closing the enormous gap between the data and the predictions of a model in which cells fire independently. Because the maximum entropy model has relatively few parameters, we evade the curse of dimensionality and associated sampling problems that would ordinarily limit the exploration of larger networks.

Key ingredients in matching theory to experiment are the focus on time windows in which the discreteness of the neural responses is evident, low spike probabilities, and weak but significant correlations among almost all pairs of cells. None of these considerations are specific to the retina. As a first test for the generality of thee ideas, we have analyzed experiments on cultured networks of cortical neurons (41), where again we find that the maximum entropy model captures over 95% of the multi–information in groups of $N = 10$ cells (Fig. 2d).

The success of a model that includes only pairwise interactions provides an enormous simplification in our description of the network. This may be important not just for us, but for the brain as well. The dominance of pairwise interactions means that learning rules based on pairwise correlations (42) could be sufficient to generate nearly optimal internal models for the distribution of "codewords" in the retinal vocabulary, thus allowing the brain to accurately evaluate new events for their degree of surprise (43).

The mapping of the maximum entropy problem to the Ising model, together with the observed level of correlations, implies that groups of $N \sim 200$ cells will behave very differently than smaller groups, and this is especially interesting because the patch of significantly correlated ganglion cells in the retina is close to this critical size (19). Because the response properties of retinal ganglion

cells adapt to the input image statistics (44; 45), this matching of correlation length and correlation strength cannot be an accident of anatomy but rather must be set by adaptive mechanisms. Perhaps there is an optimization principle which determines this operating point, maximizing coding capacity while maintaining the correlation structures which enable error–correction.

Although we have focused on networks of neurons, the same framework has the potential to describe biological networks more generally. In this view, the network is much more than the sum of its parts, but a nearly complete model can be derived from all of its pairs.

Methods

Electrophysiological Recording. Retinae from the larval tiger salamander (*Ambystoma tigrinum*) and the guinea pig (*Cavia porcellus*) were isolated from the eye retaining the pigment epithelium and placed over a multi-electrode array (19). Both were perfused with oxygenated medium: room temperature Ringers for salamander and $36°C$ Ames medium for guinea pig. Extracellular voltages were recorded by a MultiChannel Systems MEA 60 microelectrode array and streamed to disk for offline analysis. Spike waveforms were sorted either using the spike size and shape from a single electrode (19) or the full waveform on 30 electrodes (18). Recorded ganglion cells were spaced no more than $500\,\mu$m apart, and were typically close enough together to have overlapping receptive field centers. This analysis presented here is based on measurements of 95 cells recorded in 4 salamanders and 35 cells recorded in 2 Guinea pigs.

Visual Stimulation. Natural movie clips (NAT in Fig. 3) were acquired using a Canon Optura Pi video camera at 30 frames per second. Movie clips broadly sampled woodland scenes as well as man–made environments, and included several qualitatively different kinds of motion: objects moving in a scene, optic flow, and simulated saccades; see Ref. 19 for details. In spatially uniform flicker (FFF), the light intensity was randomly chosen to be black or white every 16.7 ms. For most experiments, a 20–30 s stimulus segment was repeated many times; in one experiment, a 16 min movie clip was repeated several times. All visual stimuli were displayed on an NEC FP1370 monitor and projected onto the retina using standard optics. The mean light level was 5 Lux, corresponding to photopic vision.

Cultured cortical networks. Data on cultured cortical neurons were recorded by the laboratory of S. Marom (Technion–Israel Institute of Technology) using a multi–electrode array, as described in Ref. (41). The data set analyzed here is an hour long epoch of spontaneous neuronal activity recorded through 60 electrodes.

Estimating information theoretic quantities. Information theoretic quantities such as I_N depend on the full distribution of states for the real system. Estimating these quantities can be difficult, because finite data sets lead to systematic errors (32). With large data sets (~ 1 hr) and $N < 15$ cells, however, systematic errors are small, and we can use the sample–size dependence of the estimates to correct for these errors, as in Ref. (33).

Constructing the maximum entropy distributions. For networks of modest size, as considered here, constructing the maximum entropy distribution consistent with the mean spike rates and pairwise correlations can be viewed as an optimization problem with constraints. Because the entropy is a convex function of the probabilities, and the constraints are linear, many efficient algorithms are available (46).

Acknowledgments

References

[1] Hopfield, J. J. & Tank, D. W. Computing with neural circuits: a model. *Science* **233,** 625–633 (1986).

[2] Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233,** 1416–1419 (1986).

[3] Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402 Suppl C,** 47–52 (1999).

[4] Barabási, A–L. & Oltvai, Z. N. Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics* **5,** 101–113 (2004).

[5] Perkel, D. H. & Bullock, T. H. Neural coding. *Neurosci. Res. Prog. Sum.* **3,** 221–348 (1968).

[6] Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370,** 140–143 (1994).

[7] Meister, M., Lagnado, L. & Baylor, D. A. Concerted signaling by retinal ganglion cells. *Science* **270,** 1207–1210 (1995).

[8] Riehle, A., Grun, S., Diesmann, M. & Aertsen, A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* **278,** 1950–1953 (1997).

[9] Dan, Y., Alonso, J. M., Usrey, W. M. & Reid, R. C. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat Neurosci.* **1,** 501–507 (1998).

[10] Hatsopoulos, N., Ojakangas, C., Paninski, L. & Donoghue, J. Information about movement direction obtained from synchronous activity of motor cortical neurons. *Proc Nat'l Acad Sci (USA)* **95,** 15706–15711 (1998).

[11] Abbott, L. F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput* **11,** 91–101 (1999).

[12] Bair, W., Zohary, E. & Newsome, W. T. Correlated firing in macaque visual area MT: time scales and relationship to behavior, *J Neurosci* **21,** 1676–97 (2001).

[13] Shamir, M. & Sompolinsky, H. Nonlinear population codes. *Neural Comput* **16,** 1105–1136 (2004).

[14] Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Nat'l Acad Sci (USA)* **95,** 14863–14868 (1998).

[15] Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome–wide expression data processing and modeling. *Proc Nat'l Acad Sci (USA)* **97,** 10101–10106 (2000).

[16] Holter, N. S., Maritan, A., Cieplak, M., Federoff, N. V. & Banavar, J. R. Dynamic moeling of gene expression data. *Proc Nat'l Acad Sci (USA)* **98,** 1693–1698 (2001).

[17] Meister, M., Pine, J. & Baylor, D. A. Multi–neuronal signals from the retina: acquisition and analysis. *J Neurosci Methods* **51,** 95–106 (1994).

[18] Segev, R., Goodhouse, J., Puchalla, J. L. & Berry II, M. J. Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nat Neurosci* **7,** 1155–1162 (2004).

[19] Puchalla, J. L., Schneidman, E., Harris, R. A. & Berry II, M. J. Redundancy in the population code of the retina. *Neuron* **46,** 492–504 (2005).

[20] Frechette, E. S., Sher, A., Grivich, M. I., Petrusca, D., Litke, A. M. & Chichilnisky, E. J. Fidelity of the ensemble code for visual motion in primate retina. *J Neurophyiol* **94,** 119–135 (2005).

[21] Rieke, F., Warland, D., de Ruyter van Steveninck, R. & Bialek, W. *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, 1997).

[22] Martignon, L., Deco, G., Laskey, K., Diamond, M., Freiwald, W. & Vaadia, E. Neural coding: higher–order temporal patterns in the neurostatistics of cell assemblies. *Neural Comput* **12,** 2621–2653 (2000).

[23] Grun, S., Diesmann, M. & Aertsen, A. Unitary events in multiple single-neuron spiking activity: I. Detection and significance. *Neural Comput*, **14,** 43–80 (2002).

[24] Schnitzer, M. J. & Meister, M. Multineuronal firing patterns in the signal from eye to brain. *Neuron* **37,** 499–511 (2003).

[25] Brillouin, L. *Science and Information Theory* (Academic Press, New York, 1962).

[26] Jaynes, E. T. Information theory and statistical mechanics. *Phys Rev* **106,** 62–79 (1957).

[27] Schneidman, E., Still, S., Berry II, M. J. & Bialek, W. Network information and connected correlations. *Phys Rev Lett* **91,** 238701 (2003).

[28] Landau, L. D. & Lifshitz, E. M. *Statistical Physics.* 3rd edition, revised and enlarged by EM Lifhsitz & LP Pitaevskii, translated by JB Sykes & MJ Kearsley (Pergamon, Oxford, 1978–80).

[29] Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc Nat'l Acad Sci (USA)* **79,** 2554–2558 (1982).

[30] Amit, D. J. *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, Cambridge, 1989).

[31] Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (John Wiley & Sons, New York, 1991).

[32] Estimation of entropy and information of undersampled probability distributions. NIPS 2003 Workshop, http://www.menem.com/~ilya/pages/NIPS03/.

[33] Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R. & Bialek, W. Entropy and information in neural spike trains, *Phys Rev Lett* **80,** 197–200 (1998).

[34] Lin, J. Divergence measures based on the Shannon entropy, *IEEE Trans Inf Theory* **37,** 145–151 (1991).

[35] Perkel, D. H., Gerstein, G. L. & Moore G. P. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys J* **7,** 419–440 (1967).

[36] Mezard, M., Parisi, G. & Virasoro, M. A. *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

[37] Abeles, M. *Local Cortical Circuits* (Springer–Verlag, Berlin, 1982).

[38] Aertsen, A. M., Gerstein, G. L., Habib M. K. & Palm, G. Dynamics of neuronal firing correlation: Modulation of "effective connectivity." *J Neurophysiol* **61,** 900–917 (1989).

[39] Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H. & Aertsen, A. Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* **373,** 515–518 (1995).

[40] Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D. & Yuste, R. Synfire chains and cortical songs: Temporal modules of cortical activity. *Science* **304,** 559–564 (2004).

[41] Eytan, D., Brenner, N. & Marom, S. Selective adaptation in networks of cortical neurons. *J Neurosci* **23,** 9349–9356 (2003).

[42] Bi, G. & Poo, M. M. Synaptic modification by correlated activity: Hebb's postulate revisited. *Ann Rev Neurosci* **24,** 139–166 (2001).

[43] Barlow, H. Conditions for versatile learning, Helmholtz's unconscious inference, and the task of perception. *Vision res* **30,** 1561–1571 (1990).

[44] Smirnakis, S., Berry II, M. J., Warland, D. K., Bialek, W. & Meister, M. Adaptation of retinal processing to image contrast and spatial scale. *Nature* **386,** 69–73 (1997).

[45] Hosoya, T., Baccus, S. A. & Meister, M. Dynamic predictive coding by the retina. *Nature* **436,** 71–77 (2005).

[46] Darroch, J. N. & Ratcliff, D. Generalized iterative scaling for log–linear models. *Annal Math Stat* **43,** 1470–1480 (1972).