

# A Maximum Entropy Model Applied to Spatial and Temporal Correlations from Cortical Networks *In Vitro*

Aonan Tang<sup>1\*</sup>, David Jackson<sup>4\*</sup>, Jon Hobbs,<sup>1</sup> Wei Chen,<sup>1</sup> Jodi L. Smith,<sup>3</sup> Hema Patel,<sup>3</sup> Anita Prieto,<sup>2</sup> Dumitru Petrusca,<sup>5</sup> Matthew I. Grivich,<sup>5</sup> Alexander Sher,<sup>5</sup> Pawel Hottowy,<sup>5</sup> Wladyslaw Dabrowski,<sup>6</sup> Alan M. Litke,<sup>5</sup> John M. Beggs<sup>1</sup>

Departments of <sup>1</sup>Physics and <sup>2</sup>Psychological and Brain Sciences, and <sup>3</sup>School of Medicine, Indiana University, Bloomington, Indiana 47405, <sup>4</sup>Brown University, Providence, Rhode Island 02912, <sup>5</sup>Institute for Particle Physics, University of California, Santa Cruz, California 95064, and <sup>6</sup>Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, 30-059, Krakow, Poland

Multineuron firing patterns are often observed, yet are predicted to be rare by models that assume independent firing. To explain these correlated network states, two groups recently applied a second-order maximum entropy model that used only observed firing rates and pairwise interactions as parameters (Schneidman et al., 2006; Shlens et al., 2006). Interestingly, with these minimal assumptions they predicted 90–99% of network correlations. If generally applicable, this approach could vastly simplify analyses of complex networks. However, this initial work was done largely on retinal tissue, and its applicability to cortical circuits is mostly unknown. This work also did not address the temporal evolution of correlated states. To investigate these issues, we applied the model to multielectrode data containing spontaneous spikes or local field potentials from cortical slices and cultures. The model worked slightly less well in cortex than in retina, accounting for  $88 \pm 7\%$  (mean  $\pm$  SD) of network correlations. In addition, in 8 of 13 preparations, the observed sequences of correlated states were significantly longer than predicted by concatenating states from the model. This suggested that temporal dependencies are a common feature of cortical network activity, and should be considered in future models. We found a significant relationship between strong pairwise temporal correlations and observed sequence length, suggesting that pairwise temporal correlations may allow the model to be extended into the temporal domain. We conclude that although a second-order maximum entropy model successfully predicts correlated states in cortical networks, it should be extended to account for temporal correlations observed between states.

**Key words:** microelectrode array; culture; slice; human tissue; neuronal avalanche; local field potential

## Introduction

A major hypothesis in computational neuroscience is that neurons process information through collective interactions. Despite the importance of this assumption, much is still unknown about how even small networks of neurons interact. For example, most connections between cortical neurons, as measured by correlations (Bartho et al., 2004) or synaptic strengths (Song et al., 2005), appear to be very weak, yet multineuron firing patterns are often observed in experiments. What could account for the abundance of these correlated network states?

Recent work by two groups (Schneidman et al., 2006; Shlens et

al., 2006) has shown that it is possible to predict 90–99% of network correlations in retina where interactions between neuron pairs are weak. To make these predictions, both groups used a model which considered only firing rates and pairwise interactions, and which was maximally uncommitted about all other features (a second-order maximum entropy model). This promising approach suggests that neurons do not need strong individual connections to have strong collective interactions. It further raises the possibility that activity in complex neural networks might be adequately captured with relatively simple models.

However, this new work raises several additional questions. First, how general are these results? Both groups demonstrated their models on data taken primarily from retinal tissue, although Schneidman et al. (2006) also considered one network of dissociated cortical neurons. It therefore remains unclear whether a maximum entropy approach can succeed in cortical circuitry. In addition, the models of both groups were applied to correlated states of spikes. Could such models also be applied to correlated states of local field potentials (LFPs)?

A second issue concerns the temporal evolution of correlated states. The maximum entropy model was intended to predict multineuron firing patterns at one time bin only, yet several laboratories have reported that correlated states can appear consecutively across several time bins *in vitro* (Beggs and Plenz, 2004; Ikegaya et al., 2004; Segev et al., 2004) and *in vivo* (Lindsey et al.,

Received July 24, 2007; revised Nov. 29, 2007; accepted Dec. 3, 2007.

This work was supported by National Science Foundation (NSF) Grants 0343636 (J.M.B.) and PHY-0417175 (A.M.L.), the McKnight Foundation (A.M.L.), the Burroughs Wellcome Fund Career Award at the Scientific Interface (A.S.), a MetaCyt award from the Eli Lilly Foundation to Indiana University (J.M.B.), NSF Research Experience for Undergraduates (D.J.), and a Pence foundation grant awarded to J.L.S. We are grateful to David Feldheim of the University of California Santa Cruz for use of his lab space during the 512 array recording sessions. We would also like to acknowledge Jonathon Shlens for his very helpful comments on a previous version of this manuscript and for his contributions to the spike sorting software used on the 512 electrode array data. We are also grateful to the patient and his parents who allowed us to record from his excised peritumoral cortical tissue. We thank A. Grillo, P. Grybos, and S. Kachiguine for technical development.

\*A.T. and D.J. contributed equally to this work.

Correspondence should be addressed to John M. Beggs, Indiana University Department of Physics, 727 East Third Street, Bloomington, IN 47405. E-mail: jmbeggs@indiana.edu.

DOI:10.1523/JNEUROSCI.3359-07.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/280505-14\$15.00/0

1997; Prut et al., 1998; Villa et al., 1999; Chang et al., 2000). Here we define a “sequence” as a series of correlated states appearing in consecutive time bins. Do correlated states occur in a temporally independent manner? If so, then concatenating the states predicted by the model should give a reasonable estimate of sequence lengths. However, if observed sequence lengths are longer than those produced by concatenated model states, then temporal dependencies probably exist. If these dependencies are commonly found, they should be accounted for by future models of neural network activity.

To explore the first issue, we applied a second-order maximum entropy model to a wide variety of *in vitro* cortical networks, including acute slices from rat and human tissue, as well as organotypic and dissociated cultures from rat. We explored the model’s effectiveness in predicting correlated states of both spikes and LFPs at one time bin. To address the second issue, we combined states from the model and then compared them to consecutive states observed in the actual data.

Parts of this paper (preliminary results) have been published previously in abstract form (Tang et al., 2007).

## Materials and Methods

**Tissue preparation and recording.** All neural tissue from animals was prepared according to guidelines from the National Institutes of Health and all animal procedures were approved by the Indiana University Animal Care and Use Committee. Slices of juvenile human cortex were taken only from peritumoral tissue that was removed as part of treatment for recalcitrant epilepsy. Recording from human slices was performed with the consent of the patient’s parents and was approved by the Human Subjects Committee of Indiana University.

Primary hippocampal and cortical neuronal cultures were prepared according to the methods described by (Banker and Goslin, 2002). Hippocampal and cortical tissues were dissected from Sprague Dawley rat brains (Harlan–Sprague Dawley, Indianapolis IN) of embryonic (E) day 17–18 in  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free HBSS supplemented with 10 mM HEPES, pH 7.5/1 mM sodium pyruvate (buffered Hank’s) during the dissection (all from Invitrogen, Carlsbad, CA). The tissues were incubated for 20 min in 0.25% trypsin (Invitrogen) in buffered Hanks’s at 37°C, washed 5 times with buffered Hank’s, and dissociated using fire polished Pasteur pipettes until a single cell suspension was obtained. The cells were diluted with 5 volumes of HBSS/1.5 mM  $\text{CaCl}_2$ /0.5 mM  $\text{MgCl}_2$  and centrifuged at  $200 \times g$  for 5 min. The cells were then resuspended in Neurobasal media supplemented with B-27 (Invitrogen), 0.5 mM L-glutamine (Invitrogen) and 0.025  $\mu\text{M}$  L-glutamate (Tocris Biosciences, Ellisville, MO) and 50 U/ml Penicillin and Streptomycin (Brewer et al., 1993) and seeded onto 250  $\mu\text{g}/\text{ml}$  poly-L-lysine (Sigma, St. Louis, MO) coated microelectrode arrays. The microelectrode array dishes were sealed with a membrane lid which allowed gas exchange and prevented water loss (Potter and Demarse, 2001), then placed into an incubator (NuAire, Plymouth, MN) with 5%  $\text{CO}_2$ , 9%  $\text{O}_2$  and humidified atmosphere at 37°C. Half of the media was replaced every 4 d with media having no added glutamate, effectively decreasing the concentration of glutamate by a factor of 2 every 4 d. The cells were cultured for 21 d before recording. During recording, cultures were removed from the incubator and placed in an amplifier. A pipe carrying gas from the incubator was connected to the membrane lid to prevent changes in gas content. Recordings were performed at 37°C without perfusion of medium and typically lasted 2 h.

Organotypic cultures were prepared following the method of (Stopini et al., 1991). Briefly, brains from postnatal day 1 (P1)–P3 Sprague Dawley rat pups (Harlan) were removed under a sterile hood and placed in chilled Gey’s balanced salt solution for 1 h at 8°C. After 30 min, half the solution was changed. Brains were next blocked into  $\sim 5 \text{ mm}^3$  sections containing somatosensory cortex, striatum and thalamus. Blocks were then sliced into coronal sections with a thickness of 450  $\mu\text{m}$  using a Vibratome 3000 slicer (Ted Pella, Redding, CA). Each slice was placed on a small circular cutout of permeable membrane (Millipore, Billerica, MA) that was then placed on top of a larger membrane that spanned a

culture well. Culture medium consisted of HBSS (Sigma; H9394) 1:4, Mega cell medium (Sigma; M4067) 2:4, horse serum (Sigma; H1270) 1:4, and 4 mM glutamine, with penicillin/streptomycin 1:100 volume of media (Sigma; P4083). This solution was then poured into the well so that the slices were maintained at an interface between medium below and atmosphere above. The plates of wells were next placed in an incubator and constantly maintained at 37°C in humidified atmosphere with 5%  $\text{CO}_2$ . One-half of the medium was exchanged every three d. At 3 d *in vitro*, 10  $\mu\text{l}$  of mitosis inhibitor, consisting of 4.4 mM cytosine-5- $\beta$ -arabino-furanoside, 4.4 mM uridine, and 4.4 mM 5-fluorodeoxyuridine (all from Sigma), were added for 24 h. After 3 weeks the cultures were then gently placed on a microelectrode recording array by lifting the small circular cutout of membrane with tweezers. Each culture was placed tissue side down, with the membrane facing up. We attempted to place the tissue so that neocortical layers I–V covered the array, whereas striatum and thalamus were away from the array. During recording, cultures were perfused at 1 ml/min with culture medium that was saturated with 95%  $\text{O}_2$ /5%  $\text{CO}_2$ . Recording sessions lasted 2–7 h.

Acute slices were prepared from 14- to 35-d-old Sprague Dawley rats (Harlan). Rats were deeply anesthetized with halothane and then decapitated. Brains were removed and immediately placed for 3 min in ice-cold artificial CSF (ACSF) containing (in mM) 125 sucrose, 3 KCl, 1.25  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , 26  $\text{NaHCO}_3$ , 2  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 2  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , and 10 D-glucose, saturated with 95%  $\text{O}_2$ /5%  $\text{CO}_2$ . After cooling, brains were blocked into  $\sim 5 \text{ mm}^3$  sections containing somatosensory cortex, striatum and thalamus. Blocks were then sliced into coronal sections with a thickness of 250  $\mu\text{m}$  using the tissue slicer. After cutting, slices were incubated for  $\sim 1$  h at room temperature in ACSF with the same ingredients as listed above, but with 125 mM NaCl substituted for 125 mM sucrose to restore  $\text{Na}^+$  and allow cells to fire action potentials again. After incubation, slices were adhered to microelectrode arrays with a solution of 0.1% polyethylenimine that had been previously applied and let to dry for 2 h (Wirth and Luscher, 2004). We attempted to place the tissue so that neocortical layers I–V covered the array. Slices were maintained thermostatically at 37°C and were perfused at 1.0 ml/min with ACSF solution containing 5 mM KCl and 0 mM  $\text{Mg}^{+2}$  during recording, which typically lasted 5 h. These external ionic concentrations are known to produce LFP activity in cortical brain slices (Schiff et al., 1994; Wu et al., 1999).

Acute slices of epileptogenic cortex were obtained from the brain of a 14-year-old boy with medically intractable seizures who was undergoing resection of a brain tumor and peritumoral epileptogenic zone for treatment of his seizures. To do this, a small piece of cortex was harvested by the pediatric neurosurgeon (J. Smith) from the right parietal lobe within the epileptogenic zone just posterior to the patient’s tumor. The tissue was taken from this area of cortex because it consistently demonstrated large population spikes during intraoperative electrocorticography as assessed by the surgical epileptologist (H. Patel). Immediately after its removal, the tissue was placed in a beaker of chilled, oxygenated ACSF (with sucrose substituted for NaCl, as described above) and then sliced into 250  $\mu\text{m}$  sections. Slices were then incubated for 1 h in room-temperature ACSF containing NaCl rather than sucrose. Slices were next gently placed on the 60-electrode array precoated with polyethylenimine. For recording, slices were maintained at 37°C and perfused at 1 ml/min with ACSF containing (in mM) 125 sucrose, 3.5 KCl, 1.25  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , 26  $\text{NaHCO}_3$ , 2  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 2  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , and 10 D-glucose, saturated with 95%  $\text{O}_2$ /5%  $\text{CO}_2$ . Under these conditions, LFP activity occurred spontaneously during the 1.5 h recording session.

**Electrode arrays.** All recordings, except those of spikes from organotypic cultures, were performed on microelectrode arrays from Multichannel Systems (Reutlingen, Germany). Each array had 60 electrodes, where each electrode was 30  $\mu\text{m}$  in diameter (see Fig. 1A). For organotypic cultures, electrodes were flat (see Fig. 1B), but for acute slices, electrodes came to a point 30  $\mu\text{m}$  high. In both cases, electrodes were arranged in a square grid with 200  $\mu\text{m}$  spacing between electrodes (see Fig. 1C). Spikes from organotypic cultures were recorded with a custom-made 512-electrode array (Litke et al., 2004). The flat electrodes here were 5  $\mu\text{m}$  in diameter and spaced 60  $\mu\text{m}$  apart in a hexagonal lattice (see Fig. 1B).

**Spike sorting.** Activity from dissociated cultures grown on 60-channel microelectrode arrays was sampled at 20 kHz, and waveforms that crossed a threshold of 8 SDs were cut out in 4 ms segments for later sorting (see Fig. 1A). Segments were decomposed into principal components, and sorted based on clusters identified by an operator. Putative neurons which violated a refractory period of 2 ms were excluded from analysis. Pairs of neurons sharing correlations  $\geq 0.75$  were also excluded. Sorting was done off line using software from Plexon (Dallas, TX).

Spiking activity from organotypic cortical cultures recorded on the 512-electrode array (see Fig. 1B) was sampled at 20 kHz and sorted automatically as described previously (Litke et al., 2004; Frechette et al., 2005; Shlens et al., 2007). Briefly, signals that crossed a threshold of 3 SDs were marked, and the waveforms found on the marked electrode and its six adjacent neighbors were projected into five dimensional principal component space. A mixture of Gaussians model was fit to the distribution of features based on maximum likelihood. Only the neurons that had well separated clusters in principal components space and had no refractory period violations were used in further analysis.

In cases where multiple spikes occurred in a single bin, they were treated as a single spike, as explained further below.

**LFP detection.** Extracellular activity from organotypic cultures and acute slices was sampled at 1 kHz and amplified before being stored to disk for off-line analysis. Local field potentials that showed sharp negative peaks were detected in a manner similar to that previously described (Beggs and Plenz, 2003, 2004). Briefly, if a waveform had a negative peak below a threshold set at 3 SDs of the signal, then the time of the maximum excursion was recorded as the time of that LFP (see Fig. 1C).

**Maximum entropy model.** We sought to account for the observed distribution of network states in small ensembles of neurons ( $N = 4, 6, 8, 10$ ) by using only information about firing rates and pairwise interactions. In doing this, we followed closely the methods of Schneidman et al. (2006) as well as Shlens et al. (2006) who constructed maximum entropy models of their data. Antecedents of these methods can be found in (Martignon et al., 2000; Amari, 2001; Schneidman et al., 2003), and an overview is given in (Nirenberg and Victor, 2007).

To explain this approach in more detail, it is necessary first to describe several terms. The activity of neuron  $i$  within a given time bin was represented by  $\sigma_i$ , and could take on two values: +1 for the presence of one or more spikes, and  $-1$  for the absence of spikes. Similarly, the presence (+1) or absence ( $-1$ ) of suprathreshold LFPs at electrode  $i$  was represented by  $\sigma_i$ . For brevity, we will outline the methods here for spikes, although they apply identically for suprathreshold LFPs. The expected value of  $\sigma_i$  gave information about the firing rate and was given by the following:

$$\langle \sigma_i \rangle = \frac{1}{T} \sum_{t=1}^T \sigma_i^t \quad (1)$$

where  $\sigma_i^t$  was the activity of neuron  $i$  at time  $t$ , and  $T$  is the number of time bins in the recording. The expected value of the pairwise interaction between neuron  $i$  and neuron  $j$  was given by the following:

$$\langle \sigma_i \sigma_j \rangle = \frac{1}{T} \sum_{t=1}^T \sigma_i^t \times \sigma_j^t \quad (2)$$

The state,  $V$ , of an ensemble of  $N$  neurons at a particular time was given by the following:

$$V = [\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_N]. \quad (3)$$

For  $N$  neurons, there were  $2^N$  possible states that could be observed during a recording. Not all of these states were observed equally often, and the goal of the second-order maximum entropy model was to predict the probability of observing each state, given only the firing rates  $\langle \sigma_i \rangle$ , and the pairwise interactions  $\langle \sigma_i \sigma_j \rangle$  from the data.

To predict the probability distribution of states, neural activities were mapped onto the Ising model from physics (Landau and Lifshitz, 1958; Schneidman et al., 2006) (see Fig. 2). It was assumed that each neuron

behaved like a small bar magnet, interacting with a local magnetic field and with other neighboring bar magnets. The local magnetic field around neuron  $i$  was given by  $h_i$ , and the product  $\sigma_i h_i$  gave the energy associated with this interaction. A positive value of  $\sigma_i h_i$  meant that the neuron was in harmony with its local field; this interaction was energetically favorable and more probable, as explained below. The interaction of neuron  $i$  with neuron  $j$  was given by  $J_{ij}$ , and the product  $J_{ij} \sigma_i \sigma_j$  gave the energy associated with this interaction. As with the local field, a positive value of  $J_{ij} \sigma_i \sigma_j$  meant that the neuron was in harmony with its neighbor, and this was energetically favorable and more probable. Note that when  $J_{ij}$  was positive, similar activity between neurons  $i$  and  $j$  was favored; when  $J_{ij}$  was negative, dissimilar activity was favored. As a first approximation, the local magnetic field and the interactions were given as follows:

$$h_i = \langle \sigma_i \rangle; \quad J = \langle \sigma_i \sigma_j \rangle. \quad (4)$$

The values of  $h_i$  and  $J_{ij}$  were adjusted later to provide better agreement between the model and the data, as described below.

Just as an ensemble of magnets has an energy in the Ising model, the energy  $E$  of an ensemble of  $N$  neurons in state  $V$  was given by the following:

$$E(V) = - \sum_{i=1}^N h_i \sigma_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N J_{ij} \sigma_i \sigma_j, \quad (5)$$

where summation in the second term was carried out such that  $i \neq j$ .

The next step involved mapping energies onto probabilities. To minimize unnecessary assumptions, the probability distribution with maximum entropy was chosen. The maximum entropy distribution for a data set with a given mean is an exponential distribution (Jaynes, 1957), so the probability of observing a particular state  $V_j$  was given by the following:

$$P(V_j) = \frac{e^{-E(V_j)}}{2^N}, \quad (6)$$

$$\sum_{i=1}^{2^N} e^{-E(V_i)}$$

where the denominator is just a normalizing term (the partition function in statistical mechanics) and was summed over all  $2^N$  possible states of the ensemble. Note that this equation will make states with high energy less probable than states with low energy.

Because this distribution gave the probability of each state, the expected values of the individual firing rates  $\langle \sigma_i \rangle_m$  and pairwise interactions  $\langle \sigma_i \sigma_j \rangle_m$  of the maximum entropy model could be extracted, where the subscript  $m$  denotes the following model:

$$\langle \sigma_i \rangle_m = \sum_{j=1}^{2^N} \sigma_i(V_j) \times P(V_j) \quad (7)$$

$$\langle \sigma_i \sigma_j \rangle_m = \sum_{k=1}^{2^N} \sigma_i(V_k) \times \sigma_j(V_k) \times P(V_k), \quad (8)$$

and where  $\sigma_i(V_j)$  is the activity (either +1 or  $-1$ ) of  $\sigma_i$  when the ensemble is in state  $V_j$ . The expected values from the model were then compared with  $\langle \sigma_i \rangle$  and  $\langle \sigma_i \sigma_j \rangle$  found in the data. Although the model parameters were initially selected to match the firing rates and pairwise interactions found in the data, a given set of parameters in general would not produce harmony of every neuron with its local fields and interactions for every state. To improve agreement between  $\langle \sigma_i \rangle_m$ ,  $\langle \sigma_i \sigma_j \rangle_m$  and  $\langle \sigma_i \rangle$ ,  $\langle \sigma_i \sigma_j \rangle$ , the local magnetic fields  $h_i$  and interactions  $J_{ij}$  were adjusted by an iterative scaling algorithm (Darroch and Ratcliff, 1972). Adjustments were made as follows:

$$h_i^{new} = h_i^{old} + \alpha \times \text{sign}(\langle \sigma_i \rangle) \times \log \left[ \frac{\langle \sigma_i \rangle}{\langle \sigma_i \rangle_m} \right]; \quad (9)$$

$$J_{ij}^{new} = J_{ij}^{old} + \alpha \times \text{sign}(\langle \sigma_i \sigma_j \rangle) \times \log \left[ \frac{\langle \sigma_i \sigma_j \rangle}{\langle \sigma_i \sigma_j \rangle_m} \right], \quad (10)$$



where a constant  $\alpha < 1$  was used to keep the algorithm from becoming unstable. We found  $\alpha = 0.75$  to be stable and adequately fast. Note that when the model value was less than the actual value, this produced a positive logarithm and led to an increase in either  $h_i$  or  $J_{ij}$ . Conversely, a model value that exceeded the actual value led to a negative logarithm and a reduction in  $h_i$  or  $J_{ij}$ . After adjustment, a new set of energies and probabilities was calculated for the states, and this led to new values of  $h_i$  and  $J_{ij}$ . Adjustments continued for 50,000 iterations for each ensemble, until the local field and interactions were within  $<0.1\%$  of their asymptotic values. Because the entropy is convex everywhere in this formulation, there were no local minima and methods like simulated annealing were not necessary. After iterative scaling, the final values of  $h_i$  and  $J_{ij}$  were then re-inserted into equation (5) to calculate the energy of each state, and then this energy was inserted into equation (6) to calculate the probability of observing each state  $V$ .

**Evaluating the model: one time bin.** In explaining how we evaluated the performance of the model, it is necessary to again introduce some terms. Models of several orders could be used to capture the data. A first-order model accurately represented the firing rates  $\langle\sigma_i\rangle$  found in the data, but assumed that all higher-order interactions, like  $\langle\sigma_i\sigma_j\rangle$ , were independent and were given by the product of first-order interactions:  $\langle\sigma_i\sigma_j\rangle = \langle\sigma_i\rangle \times \langle\sigma_j\rangle$ . We denoted the probability distribution produced by a first-order model by  $P_1$ , and sometimes referred to this as the “Poisson” model. A second-order model, as described above, took into account the firing rates and pairwise interactions, and produced a probability distribution that was denoted by  $P_2$ . An  $N$ th-order model would likewise take into account firing rates and all interactions from second-order (pairwise) up to  $N$ th order. Because an accurate  $N$ th-order model captured all of the higher-order interactions found in the data, its probability distribution,  $P_N$ , was identical with the probability distribution found in the data. The entropy,  $S$ , of a distribution,  $P$ , was calculated in the standard way:

$$S = - \sum_{i=1}^{2^N} P(V_i) \times \log(P(V_i)). \quad (11)$$

Because recordings were  $\sim 1$  h long, and because all ensembles were relatively small ( $N \leq 10$ ), the errors in estimating entropy were expected to be minor (Strong et al., 1998; Schneidman et al., 2006). Note that the entropy of the first-order model,  $S_1$ , was always greater than the entropy of any higher-order models,  $S_2, \dots, S_N$ , because increased interactions always reduce entropy (Cover and Thomas, 1991; Schneidman et al., 2006). The multi-information,  $I_N$ , was the total amount of entropy produced by an ensemble, and was expressed as the difference between the entropy of the first-order model and entropy of the actual data (Schneidman et al., 2003):

$$I_N = S_1 - S_N. \quad (12)$$

The amount of information accounted for by the second-order maximum entropy model was given by the following:

$$I_2 = S_1 - S_2. \quad (13)$$

The performance of the second-order maximum entropy model was therefore quantified by the fraction of the multi-information that it captured, denoted by  $f$ :

$$f = \frac{I_2}{I_N}. \quad (14)$$

This fraction could range between 0 and 1, with 1 giving perfect performance. For long recordings, the fraction can also be expressed as follows (Cover and Thomas, 1991; Shlens et al., 2006):

$$f = \frac{D_1 - D_2}{D_1}, \quad (15)$$

where  $D_1$  is the Kullback–Leibler divergence between  $P_1$  and  $P_N$  given by the following:

$$D_1 = \sum_{i=1}^{2^N} P_N(V_i) \times \log_2\left(\frac{P_N(V_i)}{P_1(V_i)}\right), \quad (16)$$

and  $D_2$  is the Kullback–Leibler divergence between  $P_2$  and  $P_N$ :

$$D_2 = \sum_{i=1}^{2^N} P_N(V_i) \times \log_2\left(\frac{P_N(V_i)}{P_2(V_i)}\right). \quad (17)$$

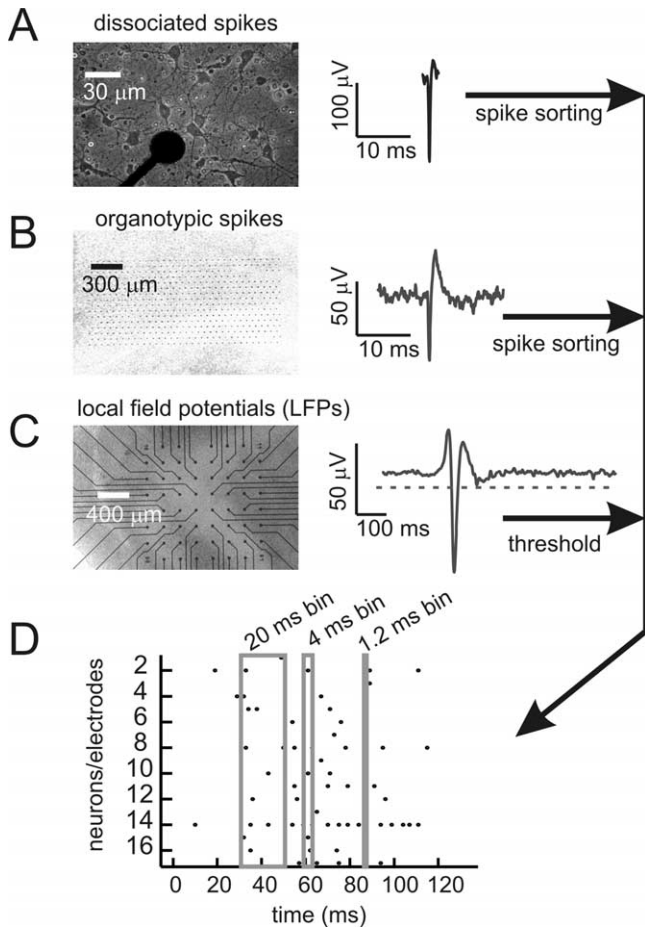
We calculated the fraction using the differences in entropy method and the Kullback–Leibler divergence method, obtaining largely similar results. For brevity, only results from the second method were reported here.

**Evaluating the model: multiple time bins.** In addition to assessing the model’s performance in predicting the probability of network states, we also sought to evaluate how well concatenated states from the model could predict sequences of active states. As before, we must first explain several terms. A sequence was defined as a series of consecutively active states, bracketed before and after by inactive states. An active state was a state in which at least one neuron was active, and an inactive state was a state in which no neuron was active. This definition of a sequence was identical to that given by Beggs and Plenz (2003) for a neuronal avalanche. Two simple statistics of sequences of states were considered: sequence length and sequence size. Sequence length was merely the number of consecutively active states, and sequence size was the total number of spikes/LFPs in the sequence. Note that this definition allowed a neuron to be counted more than once in a given sequence. We sought to examine whether the distribution of states given by the model could be used to produce a good estimate of sequence lengths and sizes observed in the data. We realized that the models of Schneidman et al. (2006) as well as Shlens et al. (2006) were not originally intended for this purpose. However, their remarkable success in predicting the distribution of states at one time point suggested to us that they might also succeed in some aspects of sequence prediction. If they did not, we reasoned, this would only point out that the appearance of correlated states was not temporally independent. This might suggest modifications to the model so that it could better predict such sequences.

To generate sequences from the data, we took the raster of activity (Fig. 1D) and randomly selected a time within its range. From this start time, we selected a period of length  $T/2$ , where  $T$  was the number of time bins in the actual recording, allowing the end of the raster to wrap around to the beginning, so that the raster was divided into two periods of equal length. From one of these half-rasters, we extracted distributions of sequence lengths and sizes. This was done for each of the 250 ensembles of neurons. We note that this procedure could have produced an artificially long sequence at the wrapping point, where the end of the raster was concatenated with the beginning. However, this occurred only once in 13 data sets and only when that data set was binned at 20 ms. This single extra sequence out of thousands did not significantly affect our conclusions. To estimate variability within the actual data group, we subtracted the 125 pairs of distributions from each other, point by point, squaring the differences and summing them. This sum of squares served as the measure of variability within the actual data group.

To generate sequences by concatenation, we drew states from  $P_2$  in proportion to their probability and then randomly concatenated them until they had a length of  $T/2$ . In this manner, an artificial raster was constructed for each of the 250 ensembles of neurons. In doing this, we assumed temporal independence between states. From these rasters, we next extracted distributions of sequence lengths and sizes. To estimate variability within this group, we randomly selected 125 pairs of distributions and subtracted them from each other, squaring the differences and summing them. This sum of squares served as the measure of variability within the concatenation group.

To measure variability between concatenated and actual data, we randomly selected 125 pairs of distributions (pair: 1 concatenated, 1 data) and subtracted them from each other, squaring the differences and summing them. This between groups sum of squares was then compared with the sum of squares from within each group. To compare these distributions, they were first sorted in ascending order and then cumulative



**Figure 1.** Data collection and representation. **A**, Left, Dissociated rat cortical neurons cultured on 60-channel microelectrode array. Several neurons can be seen around an electrode tip, which appears as a black circle. Overall array is similar to the grid shown in **C** below and has an interelectrode distance of 200  $\mu\text{m}$ . Right, Spike waveform recorded from dissociated culture. Signals that crossed a threshold were cut out and later sorted. **B**, Left, Organotypic slice culture of rat cortex on high-density array with 60  $\mu\text{m}$  interelectrode distance. Electrodes are visible as small black dots in a hexagonal lattice. Right, Spike waveform from organotypic culture. Full waveform was stored and later sorted. **C**, Left, Acute slice of human cortex removed from peritumoral region, placed on 60-channel microelectrode array. Right, LFP waveform from acute slice. All signals that crossed a threshold were recorded as events. **D**, Raster plot of data show gray boxes for the three time bins used: 20, 4, and 1.2 ms. All data were binned at 20 ms, and then also at one finer resolution (4 ms for 200  $\mu\text{m}$  array; 1.2 ms for 60  $\mu\text{m}$  array). Dots represent spikes from individual neurons or LFP negative peaks from individual electrodes. A correlated pattern occurs when events from several neurons/electrodes appear in one time bin.

probability distributions were constructed. Then the nonparametric Kolmogorov–Smirnov test was applied to probe for significant differences between the cumulative distributions. The largest difference at any point between two cumulative curves is termed  $D$ , and is the relevant statistic. If the between-groups curve significantly exceeded both of the within-groups curves, then the data were declared significantly different from the concatenated sequences.

**Significant correlations.** To assess whether a correlation was significant, we first used the hypergeometric distribution (Johnson et al., 1997) to calculate the number of times,  $M$ , that spikes from neuron A and neuron B would be expected by chance to occur in the same time bin:

$$M = \frac{(nA \times nB)}{nb}, \quad (18)$$

where  $nA$  and  $nB$  are the number of spikes produced by neurons A and B, respectively, and  $nb$  is the number of time bins in the recording. The SD of this distribution is given as follows:

$$SD = \sqrt{\frac{((nb - nA) \times (nb - nB) \times nA \times nB)}{nb^2 \times (nb - 1)}}. \quad (19)$$

For large numbers of spikes and bins, as was the case in these experiments, the hypergeometric distribution approaches a Gaussian, and we used this to estimate the 1% cutoffs in the expected number of synchronous spikes:

$$N_{+1\%} = M + 2.326 \times SD, \\ N_{-1\%} = M - 2.326 \times SD. \quad (20)$$

Here  $N_{+1\%}$  and  $N_{-1\%}$  denote the number of synchronous spikes expected in the top 1% and bottom 1%, respectively, of the distribution. With these values, we were then able to calculate the correlation values expected by chance at these 1% extremes:

$$Cov(A,A) = \frac{((nA(1 - \bar{A})^2) + ((nb - nA)\bar{A}^2))}{nb - 1} \quad (21)$$

$$Cov(B,B) = \frac{((nB \times (1 - \bar{B})^2) + ((nb - nB)\bar{B}^2))}{nb - 1} \quad (22)$$

$$Cov(A,B) = [(N(1 - \bar{A})(1 - \bar{B})) + \\ ((nA - N)(1 - \bar{A})\bar{B}) + ((nB - N)\bar{A}(1 - \bar{B})) + \\ ((nb - nA - nB + N)\bar{A}\bar{B})]/nb - 1. \quad (23)$$

$$Corr(A,B) = \frac{Cov(A,B)}{\sqrt{Cov(A,A) \times Cov(B,B)}}, \quad (24)$$

where  $\bar{A} = \frac{nA}{nb}$ , and  $\bar{B} = \frac{nB}{nb}$ , and  $N$  represents either  $N_{+1\%}$  or  $N_{-1\%}$ , depending on which extreme was being calculated. Actual correlation values that exceeded these thresholds for extreme correlations were declared significant at the  $p \leq 0.01$  level.

We calculated time-lagged correlations,  $\langle \sigma_j^t \sigma_j^{t+1} \rangle$ , as a parameter that could be used to extend the present model of spatial correlations into a model of spatiotemporal correlations. Here,  $\sigma_j^{t+1}$  represents the activity of neuron  $j$  at time  $t+1$ .

**Computing.** We wrote programs implementing the maximum entropy model in Matlab 6.5 and ran them on clusters of Power4, 1.3 GHz nodes.

## Results

The results comprise six main sections. First, we describe general features of activity in these *in vitro* cortical preparations. Second, we evaluate the performance of the second-order maximum entropy model, demonstrating that it can account for almost 90% of spatial correlations. Third, we examine the interactions ( $J_{ij}$ ) and local fields ( $h_i$ ) generated by the model, showing that these parameters are similar across diverse preparations, but vary with ensemble size. Fourth, we show that increases in group interactions with ensemble size tend to prevent the growth of entropy in cliques larger than  $\sim 40$  neurons. Fifth, we show that *in vitro* cortical networks commonly produce temporal correlations, but that these are not accounted for by concatenated model states. Sixth, we investigate how future models might account for temporal correlations.

### General description of activity

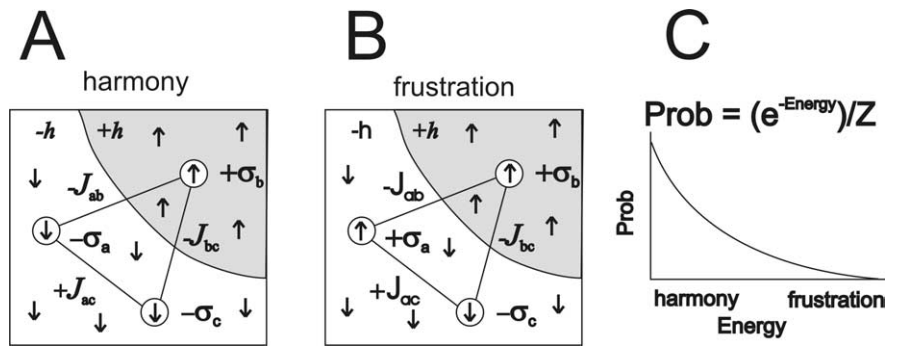
We recorded spontaneous extracellular activity from all preparations with multielectrode arrays for about 1 h ( $58.10 \pm 6.86$  min, mean  $\pm$  SD). This included spike data from dissociated cortical cultures ( $n = 3$ ) and organotypic cortex cultures ( $n = 3$ ), as well as LFP data from organotypic cortex cultures ( $n = 3$ ), rat acute cortical slices ( $n = 3$ ) and a slice of epileptogenic human cortex ( $n = 1$ ). Each preparation had 20 or more identified neurons or

active electrodes. Spiking activity from dissociated cultures and organotypic cultures consisted of periods of relative quiescence, punctuated by bursts that involved many neurons, consistent with results reported previously (Segev et al., 2002; Tateno et al., 2004; Van Pelt et al., 2004; Wagenaar et al., 2005; Baker et al., 2006; Etyan and Marom, 2006; Bettencourt et al., 2007). Spikes were biphasic waveforms  $\sim 1$  ms wide (Fig. 1A, B). Local field potential activity from organotypic cultures and acute slices was similar to that reported previously and consisted of quiescent periods also punctuated by network bursts (Beggs and Plenz, 2003; Beggs and Plenz, 2004). Local field potentials that crossed threshold appeared as negative voltage peaks  $\sim 20$  ms wide, indicative of a population spike (Fig. 1C). All data were binned at 20 ms to facilitate comparison with previous studies (Schneidman et al., 2006; Shlens et al., 2006). In addition, data collected on the 60 electrode array were also binned at 4 ms, as this was the average time between successive activation of two electrodes within a network burst when the interelectrode spacing was  $200 \mu\text{m}$ , as reported previously (Beggs and Plenz, 2003). Data from the 512 electrode array were also binned at the proportionately smaller 1.2 ms, to match the shorter interelectrode spacing of  $60 \mu\text{m}$  in that array.

When activity from all neurons or electrodes in an array was plotted in raster form, we commonly observed multiple spikes or LFPs in the same time bin (Fig. 1D). To explore the abundance of these correlated states, we randomly selected small ensembles ( $N \leq 10$ ) of neurons or electrodes from each preparation and examined the probability distributions of their states. These correlated states, or multineuron firing patterns, occurred with probabilities that were often orders of magnitude different from what was predicted by a model that assumed independent firing. A second-order maximum entropy model, analogous to the Ising model from physics (Fig. 2), was able to account for most of these correlated states. This is shown in Figure 3, where the probabilities predicted by the independent Poisson model are plotted in red dots. These red dots are often far from the black diagonal line that marks the region occupied by the actual data. The maximum entropy model, however, predicted the probability of network states far more accurately, as can be seen by the blue dots in Figure 3, which lie much closer to the diagonal line.

### Performance of the model

To quantify model performance, we calculated the fraction of ensemble correlations that was captured by the second-order maximum entropy model:  $f = I_{(2)}/I_N$ . Here,  $I_{(2)}$  was the information explained by the model and  $I_N$  was the total multi-information available in the ensemble (Fig. 4A). The figure shows that the fraction was  $>0.55$  for all 20 ms data, and generally increased toward 0.95 as the available multi-information increased. This result suggested that the model performed best when ensembles were information-rich, and was in agreement with results reported previously (Schneidman et al., 2006). The average fraction for each preparation type was also plotted for 20 ms data (Fig. 4B), and demonstrated that the model accounted



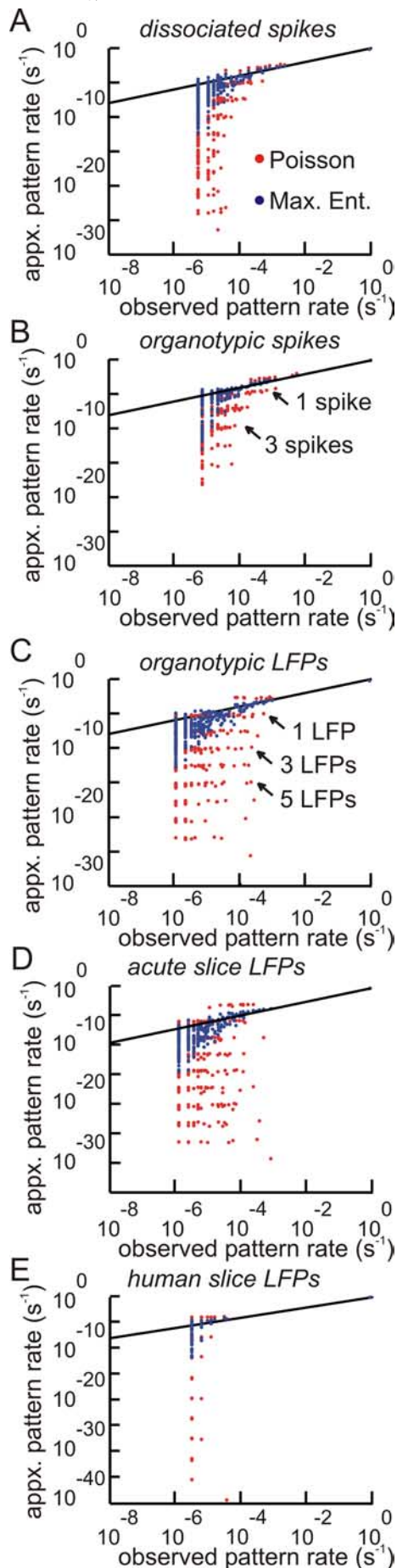
**Figure 2.** The Ising model of neural interactions. **A**, Harmony between three neurons ( $\sigma_a, \sigma_b, \sigma_c$ ), their local fields ( $-h, +h$ ), and their interactions ( $-J_{ab}, -J_{bc}, +J_{ac}$ ). Each neuron is represented by a circle, and the direction of the arrow inside depicts whether the neuron is firing a spike (up) or not (down). The gray region surrounding neuron  $\sigma_b$  has many arrows pointing up, indicating the positive local field  $+h$ . The white region surrounding neurons  $\sigma_a$  and  $\sigma_c$  has many downward pointing arrows, indicating the negative local field  $-h$ . When the arrow of a neuron is pointing in the same direction as its local field, it is in harmony with that field. Neurons in **A** are also in harmony with their interactions, represented by the solid black lines joining the neurons. A positive interaction like  $+J_{ac}$  tends to make neurons share the same state (both up, or both down), whereas a negative interaction like  $-J_{ab}$  tends to make neurons have different states (one up, one down). **B**, When neuron  $\sigma_a$  changes its state and fires a spike, it introduces frustration. Now  $\sigma_a$  has an arrow pointing against its local field and against its interactions with neurons  $\sigma_b$  and  $\sigma_c$ . To denote this frustration, the parameters  $-h, -J_{ab}$ , and  $+J_{ac}$  are displayed in a different font. **C**, Harmonious states in the Ising model are lower in energy and more probable than frustrated states. The exponential probability distribution maximizes entropy, given the constraints of the local fields and the interactions. The  $Z$  in the equation just ensures that the probability distribution sums to one.

for a sizable proportion of ensemble correlations in widely different preparations including dissociated cultures, organotypic cultures and acute slices (supplemental Table 1, available at www.jneurosci.org as supplemental material). The average fraction for all preparations was  $0.88 \pm 0.07$ , in agreement with the 0.90 fraction reported by Schneidman et al. (2006), but below the 0.98 value reported by Shlens et al. (2006). Figure 4, C and D, presents the same results for data binned at shorter times (1.2 or 4 ms). The model performed better on data binned at 20 ms ( $f = 0.88 \pm 0.07$ ) than on data binned at finer temporal resolutions ( $f = 0.76 \pm 0.12$ ). This difference was significant (Mann–Whitney two-tailed test,  $n_1 = 13; n_2 = 13; U = 138; p < 0.006$ ). Interestingly, the model also performed significantly better on LFP data ( $f = 0.93 \pm 0.04; n = 7$  preparations; 1750 ensembles; 20 ms data) than on spike data ( $f = 0.85 \pm 0.06; n = 6$  preparations; 1500 ensembles; 20 ms data) (Mann–Whitney two-tailed test,  $n_1 = 6; n_2 = 7; U = 41; p < 0.004$ ), suggesting that the maximum entropy approach is applicable to neural data that is not just in the form of spikes.

### Interactions $J$ and local fields $h$

To further compare the cortical network data with previously published results from mostly retinal tissue (Schneidman et al., 2006), we examined the values of the interactions  $J_{ij}$  and local fields  $h_i$  produced by the model. Figure 5 shows representative results from one set of spike data and one set of LFP data. In general, the results for the cortical organotypic spike data were similar to those published for retinal spike data (Schneidman et al., 2006). The interaction values  $J_{ij}$  were both positive and negative, but became increasingly positive for larger pairwise correlation coefficients  $C_{ij}$  (Fig. 5A, left). This suggested that weak interactions (low  $C_{ij}$ ) could be either excitatory (positive  $J_{ij}$ ) or inhibitory (negative  $J_{ij}$ ), but that strong interactions were dominated by excitatory connections. The distribution of local fields  $h_i$  had a negative mean (Fig. 5B, left), suggesting that most neurons were biased to not fire spikes. The distribution of interactions  $J_{ij}$  had a slightly positive mean but was near zero (Fig. 5C, left), suggesting that the average interaction between neurons was pos-





itive but weak. A comparison of the average interactions  $J_{ij}$  for LFP and spike data binned at 20 ms revealed that they were not significantly different ( $J_{\text{LFP}} = 0.73 \pm 0.08$ ;  $J_{\text{Spike}} = 0.66 \pm 0.04$ ; Mann–Whitney two-tailed test,  $n_1 = 6$ ,  $n_2 = 7$ ,  $U = 33$ ,  $p > 0.05$ ), although individual distributions sometimes appeared to be dissimilar. In addition, the pairwise correlation coefficients for LFP data were not significantly larger than those for spike data binned at 20 ms ( $C_{\text{LFP}} = 0.38 \pm 0.26$ ;  $C_{\text{Spike}} = 0.13 \pm 0.12$ ; Mann–Whitney two-tailed test,  $n_1 = 6$ ,  $n_2 = 7$ ,  $U = 33$ ,  $p > 0.05$ ), indicating that the data sets were similar in this respect.

The fact that both positive and negative interactions  $J_{ij}$  coexisted in the same cortical networks meant that many ensembles were “frustrated” and could not attain harmony in any state (Fig. 2B). We measured the number of frustrated triplets (triads of neurons or LFPs that had an odd number of negative interactions) and found that frustration was indeed common. For LFP data, the fraction of frustrated triplets was  $(0.46 \pm 0.05)$ , and for spike data the fraction was  $(0.43 \pm 0.05)$  (Table 1). Frustration was prevalent in data binned at 20 ms and at 1.2 or 4 ms. Networks with no frustration possess one harmonious state and are predicted to have little diversity, whereas networks with frustration possess many equally likely states and are predicted to have a wider range of activity patterns.

We also examined interactions  $J_{ij}$  and local fields  $h_i$  produced by the model as a function of ensemble size. Figure 6A shows that interactions for four-cell ensembles tended to be larger than interactions for 10-cell ensembles. This is seen by noting that more dots were below the diagonal line than above it. Although this was particularly noticeable for the spike data (left panel), it was also true for the LFP data (right panel). In fact, as ensemble size was increased from four to 10, there was a significant decrease in interaction strength for data binned at 20 ms (ANOVA,  $df = 3, 39$ ;  $p < 0.05$ ) (Table 2). This was also true for data at short (1.2 or 4 ms) bin widths (ANOVA,  $df = 3, 39$ ;  $p < 0.05$ ) (Table 3). Our results are in contrast to those reported by Schneidman et al. (2006), suggesting that cortical networks behave differently from retinal networks in this respect. In addition, the local field  $h_i$  generated by intrinsic neuronal activity became significantly less negative with increasing ensemble size for data binned at long and short bin widths (ANOVA,  $df = 3, 39$ ;  $p < 0.05$ ) (Tables 2, 3), consistent with previously reported results (Schneidman et al., 2006). Although this seemed to imply that neurons in larger ensembles would be less constrained by their neighbors, this was not the case. Following Schneidman et al. (2006), we examined the local field produced by interactions with other cells,  $h_i^{\text{int}}$ , where

$$h_i^{\text{int}} = \frac{1}{2} \sum_{j \neq i} J_{ij} \sigma_j.$$

Whereas  $h_i$  tended toward zero with larger ensemble sizes,  $h_i^{\text{int}}$  became significantly more negative for data at short, but not at

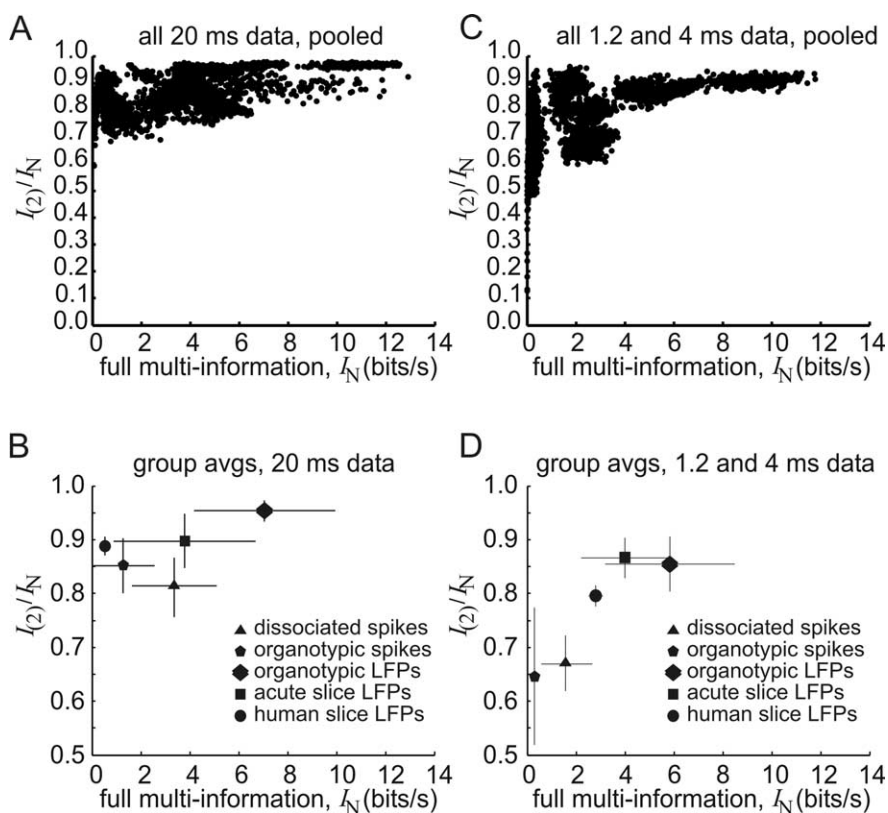
←

**Figure 3.** The maximum entropy model successfully predicts correlated states for many *in vitro* preparations. Figures compare the abundance of correlated states from data and models. The predicted rate is plotted against the observed rate. Correlated states from actual data would lie along the diagonal line. Red dots represent rates for 250 ensembles of 10 cells, each predicted by an independent Poisson model. Blue dots represent rates predicted by the second-order maximum entropy model. Note how blue dots in general lie much closer to the diagonal line, indicating superior prediction over the Poisson model. Horizontal bands of red dots represent multievent states predicted by the Poisson model, and often lie over 10 orders of magnitude below diagonal line. **A–E** show results on a single representative preparation of each type. For each preparation, 250 ensembles each containing 10 randomly chosen cells were run. Results are shown for data binned at 20 ms. Data binned at finer resolutions fit less well.

long, bin widths (ANOVA,  $df = 3, 39; p > 0.05$ ) (Tables 2, 3), in partial agreement with previous work (Schneidman et al., 2006). The increased negative strength of  $h^{int}$  can be seen in Figure 6B, where the grayscale density cloud for 4 cells (top) tended to move further across the diagonal line for 10 cells (bottom). This movement reflected the trend for  $h^{int}$  to increase with ensemble size, indicating that in large ensembles single-neuron activity was quite strongly influenced by the activity of neighboring neurons. This can also be seen in Figure 6C, where both  $h_i$  and  $J_{ij}$  tended to zero with increasing ensemble size, whereas  $h^{int}$  became increasingly negative. Opposing trends in  $h^{int}$  and  $h_i$  are to be expected, although, because the sum  $h^{int} + h_i$  should roughly equal the constant value  $\langle \sigma_i \rangle$ . The increasingly strong influence of neuronal neighbors suggested that larger ensembles would become constrained in the number of states that they visited.

### Neuronal clique sizes

To estimate how neuronal interactions would constrain larger ensembles, we plotted the multi-information,  $I_N$ , and the entropy produced by the first-order model,  $S_1$ , together in log–log coordinates as ensemble size was increased (Schneidman et al., 2006). Because  $S_N = S_1 - I_N$ , we were able to estimate the actual entropy in the ensemble,  $S_N$ , as the difference between these two lines. This estimate assumed that both  $S_1$  and  $I_N$  could be extrapolated by a linear function in log–log space. As seen in Figure 7, these lines converged in all data sets, meaning that  $S_N$  would be expected to stop growing for ensembles beyond a particular size. For example, the bold white circle in Figure 7A occurs between 20 and 30 cells, indicating that the entropy produced by ensembles from dissociated cultures containing  $\sim 25$  neurons would be nearly maximal. Ensembles larger than  $\sim 25$  cells would not be expected to have larger entropy, although they would have more states that could be visited. Here, we use the term “clique” to describe the ensemble size at which entropy would stop growing. The average clique size for data binned at 20 ms was  $23.0 \pm 27.3$ , and was  $38.7 \pm 36.9$  for data binned at shorter intervals (supplemental Table 2, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). These numbers are smaller than those reported by Schneidman et al. (2006), who reported clique sizes of  $\sim 180$  neurons in retinal tissue. Interestingly, in our data the clique sizes for spike data were significantly larger than for LFP data at both bin widths of data (Fig. 7F) (20 ms spikes,  $39.6 \pm 33.8$ ; 20 ms LFPs,  $8.8 \pm 4.5$ ; Mann–Whitney two-tailed test,  $n_1 = 6, n_2 = 7, U = 40, p < 0.005$ ; 1.2/4 ms spikes,  $65.4 \pm 38.9$ ; 4 ms LFPs,  $15.7 \pm 11.8$ ; Mann–Whitney two-tailed test,  $n_1 = 6, n_2 = 7, U = 39, p < 0.005$ ). Spike cliques were larger by a factor of 4.5 for data binned at 20 ms, and by a factor of 4.2 for data binned at 1.2 or 4 ms. This nearly constant ratio is consistent with the idea that LFP population spikes are produced by groups of several neurons firing action potentials synchronously near the electrode.



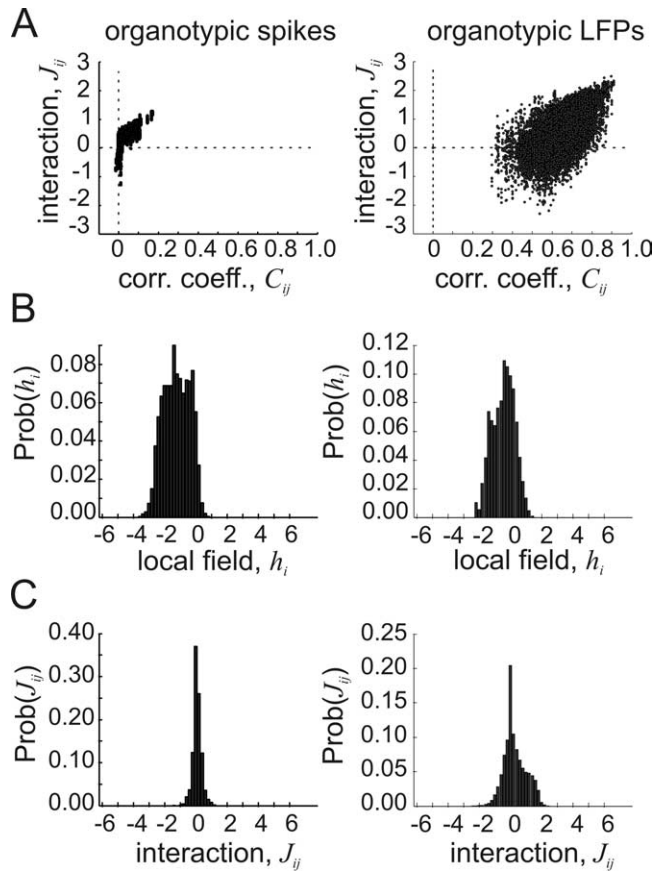
**Figure 4.** Quantifying performance in the maximum entropy model. **A**, The fraction of network correlations captured by the model,  $I_{(2)}/I_N$ , is plotted against the full multi-information  $I_N$ , for data binned at 20 ms. Each dot represents one ensemble of 10 cells, randomly chosen from one preparation. Each of the 13 preparations had 250 ensembles, giving 3250 dots. The general trend is for fractions to approach 0.95 as multi-information increases. **B**, Average fraction of network correlations for each preparation type, for data binned at 20 ms. Each symbol is an average from three preparations of that type, except for the human slice, where there was only one preparation. **C**, The fraction of network correlations captured by the model for data binned at 1.2 and 4 ms. **D**, Average fraction of network correlations for each preparation type, for data binned at 1.2 and 4 ms. Error bars show SDs. Note that model performance was better for LFP data and better for 20 ms data.

### Temporal correlations not captured by concatenating states

The above results indicated that the second-order maximum entropy model was effective in predicting correlated states for cortical tissue. These correlated states could be thought of as snapshots of the spatial correlation structure. We also sought to predict sequences of snapshots, like frames in a short movie segment, by concatenating states from the model. Although the model was not originally intended to address temporal correlations, we chose to examine whether temporal correlations were abundant in the data. If so, we reasoned, then they should be accounted for in future models.

As a simple measure of temporal structure, we plotted the distributions of sequence lengths and sizes for actual data (Fig. 8) and compared them to the distributions produced by concatenating states drawn from the model. Here, we defined a sequence as a series of consecutively active states, bracketed at the beginning and end by inactive states (see Materials and Methods). In the actual data, there was a strong tendency for one active state to be followed by another active state, often producing sequences longer than one frame. However, the sequences produced by concatenating states from the model were often very short because the most common state in the model distribution was the inactive state. This difference can be seen in the log–log plots in the left column of Figure 8, where the average length distributions produced by actual data (black dots) had tails that were often far longer than the average length distributions produced by concat-



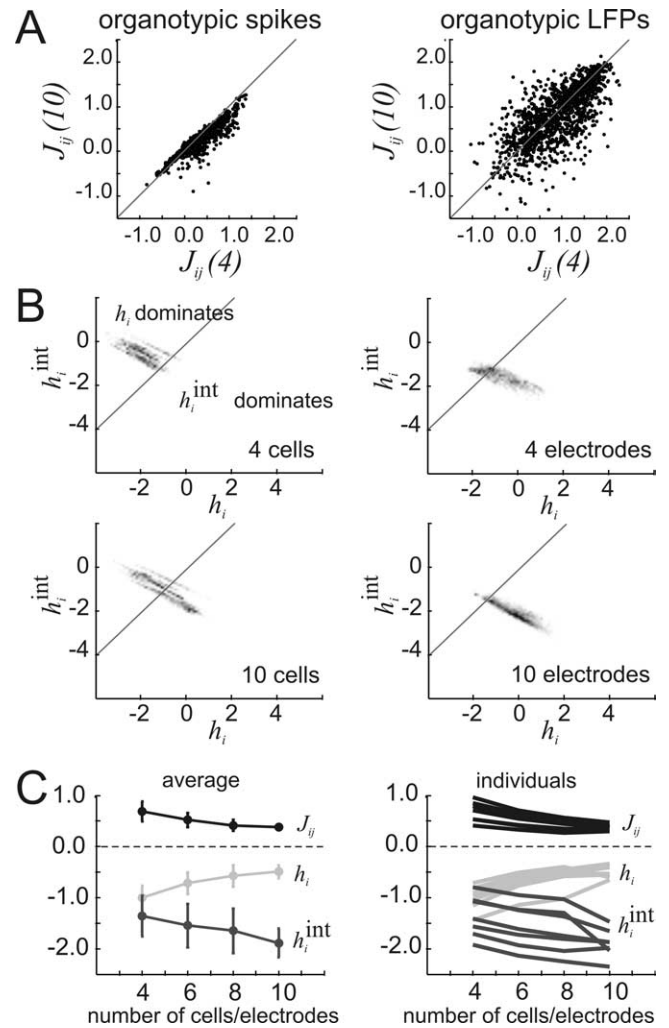


**Figure 5.** Interactions and local fields for representative preparations with spikes and with LFPs. **A**, Left, Interaction strengths  $J_{ij}$  plotted against correlation coefficients  $C_{ij}$  for spike data from an organotypic culture. Each dot represents one neuron pair from one ensemble; 2500 pairs are shown. Right, Interaction strengths plotted against correlation coefficients for LFPs from an organotypic culture. Note much larger correlation coefficients and wider range of interaction strengths in LFP data. **B**, Left, Distribution of local field strength  $h_i$  for organotypic spikes. The distribution mean is negative, indicating that most cells are influenced to not spike by other cells outside the ensemble. Right, Distribution of field strength for organotypic LFPs shows similar negative mean. **C**, Distribution of interaction strengths for organotypic spikes, left, and for organotypic LFPs, right. Positive and negative interactions exist in both preparations, suggesting that frustration is common (Table 1). Results are shown for data binned at 20 ms.

**Table 1. Fraction of frustrated triplets**

File type	20 ms bin width	1.2/4 ms bin width
Dissociated spikes	0.481	0.464
	0.461	0.252
	0.476	0.472
Organotypic spikes	0.377	0.025
	0.396	0.014
	0.377	0.001
Organotypic LFPs	0.467	0.452
	0.488	0.496
	0.490	0.477
Acute LFPs	0.350	0.390
	0.502	0.468
	0.493	0.475
Human LFPs	0.430	0.452

The fraction of frustrated triplets for all networks at two different bin widths. A frustrated triplet consists of three neurons or electrodes with conflicting interaction strengths  $J_{ij}$ . These conflicts prevent any state of the triplet from achieving harmony. Note that frustration is common in most preparations, suggesting that it is difficult for networks to settle into states with long-term stability.



**Figure 6.** Interactions,  $J_{ij}$  and local fields,  $h_i$ , change with ensemble size. **A**, Left, Interaction strengths for cell pairs chosen from a 10-cell ensemble,  $J_{ij}^{(10)}$ , plotted against interaction strengths for the same cell pairs chosen from a 4-cell ensemble,  $J_{ij}^{(4)}$ , for representative organotypic spike data. Right, Same plot now with representative organotypic LFP data. Note that asymmetry in spike plot is not present in LFP plot, indicating that in this case interaction strength decreases with ensemble size more for spikes than for LFPs. **B**, Plots of local field strength caused by interactions,  $h_i^{\text{int}}$ , against local field strength caused by intrinsic activity,  $h_i$ . Density cloud moves across the diagonal line as ensemble size is increased from four to 10, indicating that  $h_i^{\text{int}}$  dominates over  $h_i$  in larger ensembles. Higher density is coded by darker pixels. **C**, Left, Average of  $J_{ij}^{(10)}$ ,  $h_i$ , and  $h_i^{\text{int}}$  for all preparations as ensemble size is increased. Significant changes with ensemble size occurred for  $J_{ij}$  ( $F = 5.91$ ,  $F_{\text{crit}} = 2.85$ ,  $\alpha = 0.05$ ,  $df = 3, 39$ ) and  $h_i$  ( $F = 7.65$ ), but not for  $h_i^{\text{int}}$  ( $F = 2.54$ ). Right, Individual plots of  $J_{ij}^{(10)}$ ,  $h_i$ , and  $h_i^{\text{int}}$  for each preparation as ensemble size is increased, showing that average trends were followed in each individual case. Error bars give SDs. Results shown are for data binned at 20 ms. Full results are given in Tables 2 and 3.

enation (gray dots). We also compared sequence sizes produced by the data and concatenation (Fig. 8, right column). To quantify differences between the data and concatenation, we used resampling to generate many data and concatenation distributions, and then compared them using the nonparametric Kolmogorov–Smirnov test (see Materials and Methods) (Fig. 9). In 8 of 13 data sets binned at 20 ms, the sequence lengths produced by the data were significantly longer than those produced by concatenations of model states (supplemental Table 3, available at www.jneurosci.org as supplemental material). This trend was even stronger for data binned at 1.2 or 4 ms, where 11 of 13 data sets had significantly longer sequence length distributions (supple-

**Table 2. Trends in  $J_{avg}$ ,  $h_{avg}$ , and  $h_{avg}^{int}$  (20 ms bin)**

Data type	Variable	$F_{actual}$	$F_{crit}$	df
All data, pooled	$J_{avg}$	5.91*	2.85	3,39
	$h_{avg}$	7.65*		
	$h_{avg}^{int}$	2.54		
Spike data only	$J_{avg}$	2.46	3.16	3,18
	$h_{avg}$	2.94		
	$h_{avg}^{int}$	1.54		
LFP data only	$J_{avg}$	7.31*	3.07	3,21
	$h_{avg}$	8.92*		
	$h_{avg}^{int}$	2.23		

The interactions  $J_{avg}$  and local fields  $h_{avg}$ ,  $h_{avg}^{int}$  change in strength as ensemble size is increased from 4 to 10. Interactions ( $J_{avg}$ ) and local fields caused by cells outside the ensemble and intrinsic firing tendencies ( $h_{avg}$ ) moved toward zero with increasing ensemble size (Fig. 6). However, local fields caused by cells within the ensemble ( $h_{avg}^{int}$ ) became more negative. Significant changes at the  $p < 0.05$  level are indicated by an asterisk. Table shows only data binned at 20 ms.

**Table 3. Trends in  $J_{avg}$ ,  $h_{avg}$ , and  $h_{avg}^{int}$  (1.2 and 4 ms bins)**

Data type	Variable	$F_{actual}$	$F_{crit}$	df
All data, pooled	$J_{avg}$	47.31*	2.85	3,39
	$h_{avg}$	47.68*		
	$h_{avg}^{int}$	18.72*		
Spike data only	$J_{avg}$	15.07*	3.16	3,18
	$h_{avg}$	12.46*		
	$h_{avg}^{int}$	9.16*		
LFP data only	$J_{avg}$	58.07*	3.07	3,21
	$h_{avg}$	70.21*		
	$h_{avg}^{int}$	10.53*		

The interactions  $J_{avg}$  and local fields  $h_{avg}$  and  $h_{avg}^{int}$  change in strength as ensemble size is increased from 4 to 10. All changes were significant at the  $p < 0.05$  level and are indicated by an asterisk. Table shows only data binned at 1.2 or 4 ms.

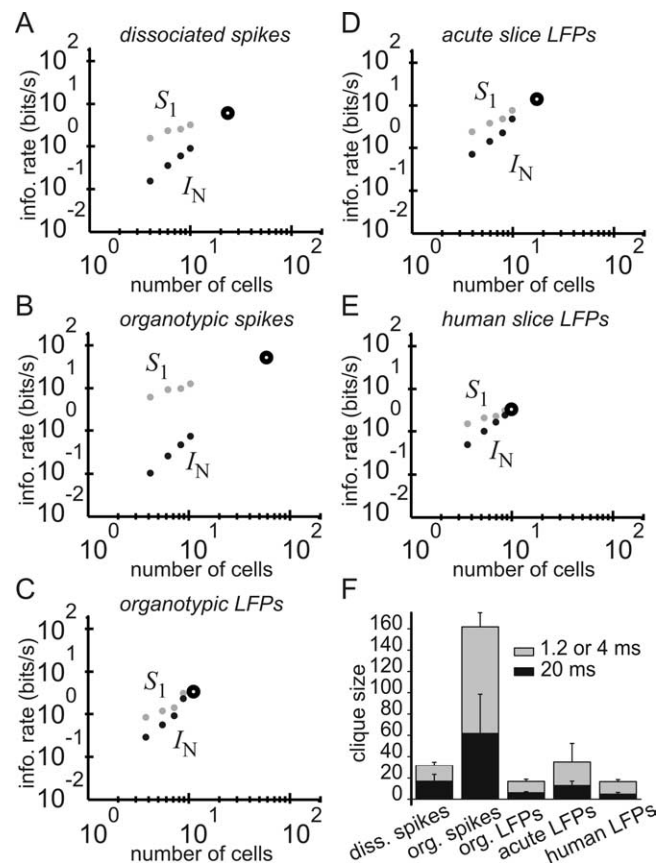
mental Table 3, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). In contrast, the sequence size distributions (size as the total number of spikes or LFPs in a sequence) were significantly larger in only 3 of 13 data sets binned at 20 ms, and in 4 of 13 data sets binned at 1.2 or 4 ms (supplemental Table 3, available at [www.jneurosci.org](http://www.jneurosci.org) as supplemental material). Overall, these results indicate that sequences of active states do not occur independently in cortical networks *in vitro*.

### Possible temporal extensions of the model

Parameters that could account for delayed relationships between neurons might allow the present model to be extended into the temporal domain (Fig. 10A). We considered the time-lagged correlation,  $\langle \sigma_i^t \sigma_j^{t+1} \rangle$ , as a candidate for this task. We hypothesized that large time-lagged correlations would reflect strong synaptic connections between neurons over time. Therefore we expected that strong time-lagged correlations would be significantly related to long sequence lengths. To measure how much the sequence lengths in the actual data exceeded the sequence lengths produced by concatenation, we used the Kolmogorov–Smirnov statistic  $D$ . The fraction of statistically significant time-lagged correlations in data binned at 20 ms indeed had a significant relationship with  $D$ , as shown in Figure 10B ( $r = 0.627$ ;  $p = 0.022$ ). This relationship was even stronger for data binned at finer temporal resolutions (Fig. 10C) ( $r = 0.735$ ;  $p = 0.004$ ). These findings suggest that the time-lagged correlation would be a plausible parameter to include in a future model that would address temporal relationships between correlated network states.

### Discussion

There are two main findings of this work. First, it demonstrates that a second-order maximum entropy model predicts correlated

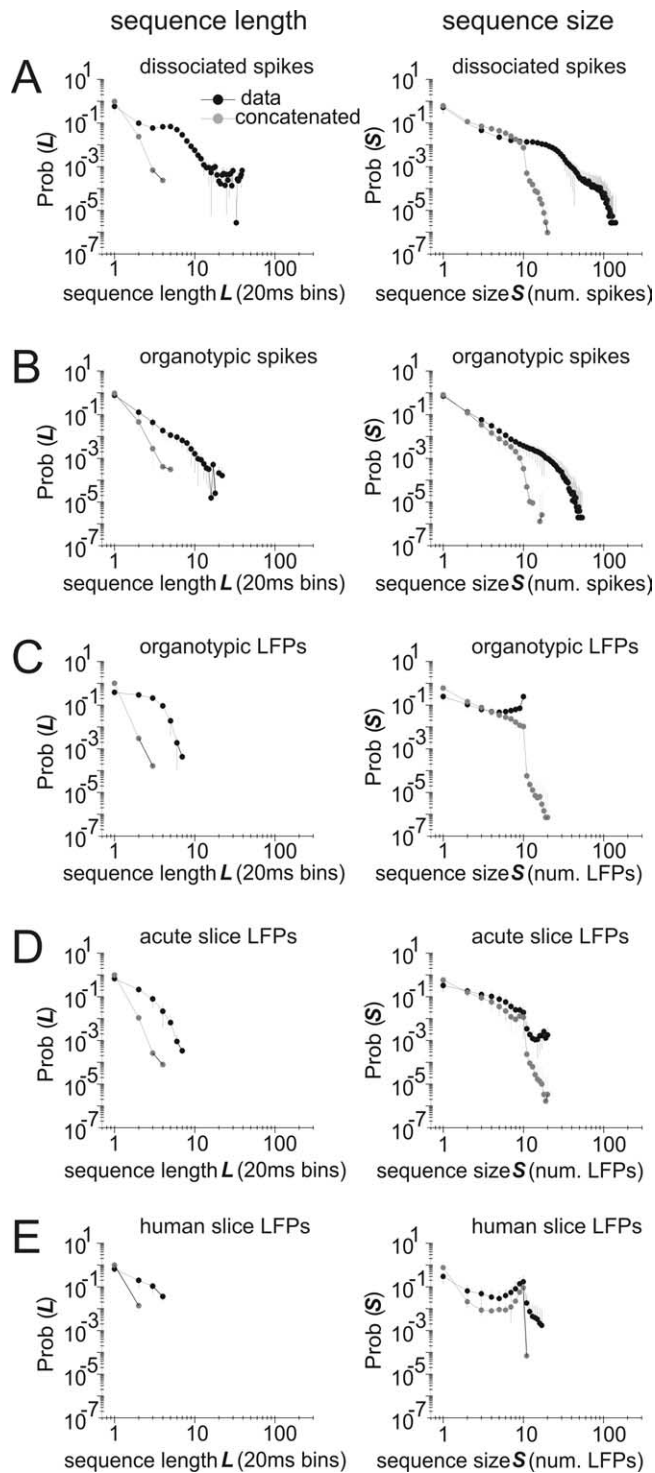


**Figure 7.** Estimation of neuronal clique sizes. Information rate for independent model ( $S_1$ , gray dots) and for actual data ( $I_N$ , black dots) are plotted in log–log space as a function of number of cells in ensemble. The point at which the  $S_1$  and the  $I_N$  extrapolated lines would intersect gives an estimate of the ensemble size at which entropy would cease to increase. **A–E**, Representative plots for dissociated culture data, organotypic spike data, organotypic LFP data, acute slice LFP data, and human slice LFP data binned at 20 ms. **F**, Histogram showing average clique sizes for each data type, for both short (1.2 or 4 ms, gray) and long (20 ms, black) bin widths. Error bars give SDs. Results suggest that entropy will no longer increase with ensemble size in relatively small cliques.

states for cortical networks *in vitro* with reasonable accuracy, for both spike and LFP data. Second, it demonstrates that correlated states have significant temporal dependencies, and that these cannot be captured by concatenating states from the model. In what follows, we will discuss the validity of this work, how it compares to previous studies, and the implications it has for dynamic models of neural networks.

### Validity

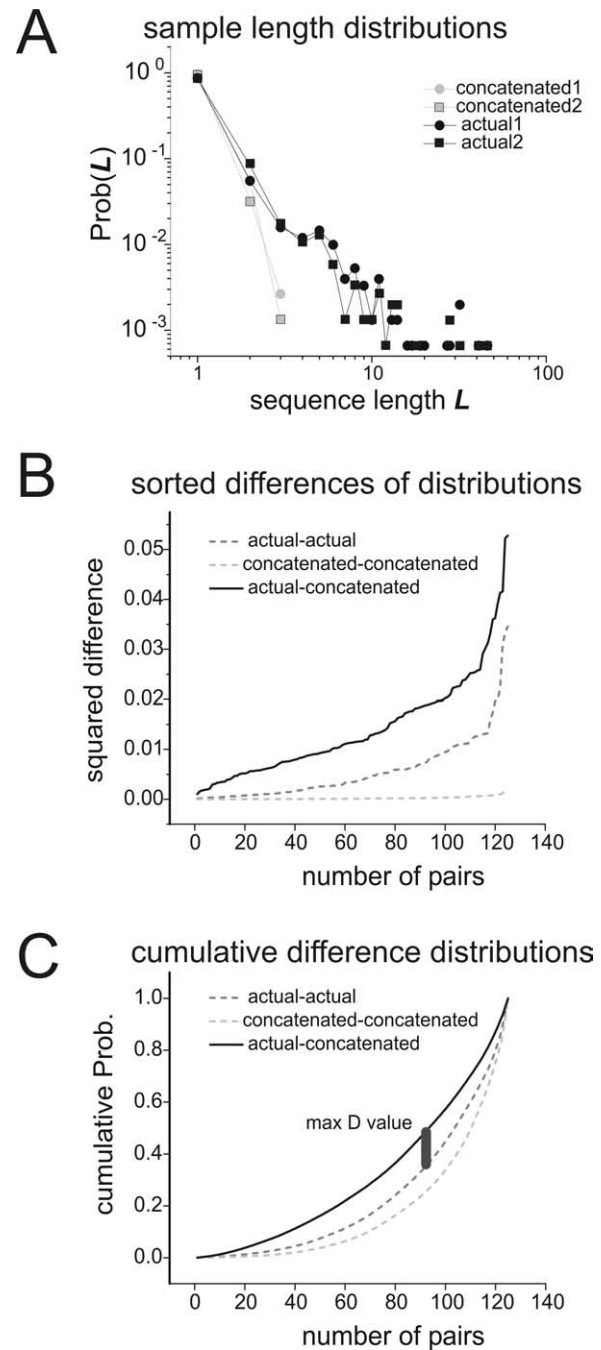
Because most of the previous work with the maximum entropy approach (Schneidman et al., 2006; Shlens et al., 2006) was done in retinal tissue, it is not easy to directly assess the validity of the present results. However, Schneidman et al. (2006) also applied the model to one dissociated culture of cortical neurons. They reported that the model was able to account for >90% of the available multi-information in one dissociated culture, whereas we obtained a fraction of 79–85% in three of the same preparation type. Although this appears to be a discrepancy, other general features reported by Schneidman et al. (2006) were mirrored by our findings: the model performed better on information-rich ensembles;  $h^{int}$  increased with ensemble size;  $h$  tended toward zero with increased ensemble size; larger ensembles were predicted to form cliques where entropy would cease to increase; the



**Figure 8.** Sequence lengths and sizes predicted by concatenation are often significantly smaller than those found in actual data. Concatenated sequences were constructed by randomly drawing states from the model probability distribution. The left column shows sequence length distributions for concatenated sequences (gray) and data (black); the right column shows sequence size distributions. *A–E* show representative plots for data from dissociated spikes, organotypic spikes, organotypic LFPs, acute slice LFPs, and human slice LFPs. Error bars show SDs. Results shown are for data binned at 20 ms.

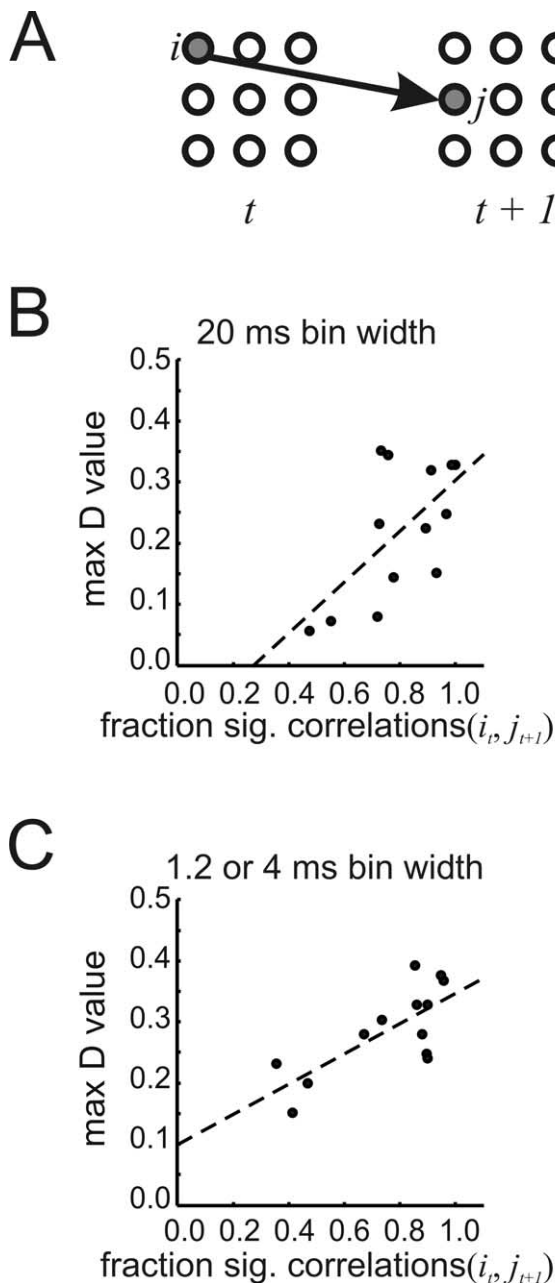
fraction of frustrated triplets was  $\sim 40\%$ . All of these findings suggest that our implementation of the maximum entropy model was quite similar to that of Schneidman et al. (2006).

A major conclusion from maximum entropy approaches is



**Figure 9.** Comparing sequence length distributions. *A*, Two sample sequence length distributions from the data and from concatenations, plotted together in log–log coordinates. Sample distributions from the data are plotted in black; those from concatenations are in gray. Sample distributions from the data were taken from a half-length segment of the recording with a randomly chosen start time. To construct distributions from concatenation, correlated states were chosen in proportion to their probability of occurrence in the model and then concatenated. Note the long tail in the data distributions. *B*, Sum of squares differences between length distributions are sorted in ascending order and plotted. The light gray dashed curve is for differences between pairs of concatenated length distributions; the dark gray dashed curve is for differences between pairs of data length distributions; the solid black curve is for differences between pairs of data and concatenated length distributions. For each curve, 125 pairs of sample distributions like those shown in *A* were used. Note that differences between concatenations and data are larger than differences within each group. *C*, Cumulative distributions of differences. The curves in *B* were converted to cumulative probability distributions, and the maximum difference between the (data–concatenated) curve and the (data–data) curve was found. This difference was called  $D$  and was used to test for significance in the Kolmogorov–Smirnov test. Thus,  $D$  was a measure of the temporal mismatch between concatenations and data.





**Figure 10.** Candidate factor for a temporal model. **A**, Schematic showing temporal relationship of activity in neurons/electrodes  $i$  and  $j$  across time. The appropriate relationship might allow the maximum entropy model to be extended to account for transitions between correlated states. **B**, The fraction of significant temporal correlations is significantly correlated with maximum  $D$  value for 20 ms data ( $r = 0.627$ ;  $p = 0.022$ ). **C**, The fraction of significant temporal correlations is also significantly correlated with maximum  $D$  value in 1.2 or 4 ms data ( $r = 0.735$ ;  $p = 0.004$ ). To calculate the fraction of significant correlations, all correlations between neuron/electrode pairs across one time step were compared with correlations expected by chance.

that higher-order correlations play a relatively minor role in living neural networks. It is therefore reasonable to question whether such approaches can detect higher-order correlations if they were present. This issue has been addressed previously, so we did not undertake a sensitivity analysis ourselves. However, Shlens et al. (2006) showed that when higher-order correlations were injected into actual data, the performance of the maximum entropy model declined. Because our approach is mathematically similar, we conclude that we would be able to detect higher-order

correlations in our data as a reduction in the fraction of multi-information captured by the second-order model.

Although the *in vitro* preparations used here are not driven by sensory inputs and are disconnected from other brain regions, they have several features that make them popular for investigation. Simplified model systems are typically easier to understand than complex systems in their entirety, yet often retain important emergent properties. For example, many network phenomena observed in the intact organism can also be seen *in vitro* [e.g., gamma and theta oscillations (Traub et al., 1996; Fischer et al., 2002), slow wave oscillations (Sanches-Vives and McCormick, 2000), up-states (Cossart et al., 2003), reproducible activity patterns (Ikegaya et al., 2004; Beggs and Plenz, 2004), and neuronal avalanches (Beggs and Plenz, 2003; Petermann et al., 2006)]. In addition, *in vitro* preparations allow favorable recording conditions like short interelectrode distances (200–60  $\mu\text{m}$ ), large numbers of electrodes (60–512) and long-term stability (1–10 h). Such features improve the chances of recording from synaptically connected neurons and enhance the statistical power of the data set. For these and other reasons, *in vitro* approaches will probably continue to provide important information about network function, and play an essential role in neuroscience research that is complementary to *in vivo* work. Nevertheless, caution should be used in extrapolating *in vitro* results to *in vivo* systems.

The three acute cortical slices used in this study were bathed in fluid containing elevated potassium (5 mM) and reduced magnesium (0 mM). These ionic concentrations are different from what would typically be found *in vivo*. Therefore we also chose to examine activity in many other cortical network preparations bathed in culture medium or in normal CSF. In addition, the one slice of human tissue that we studied was removed from a patient with epilepsy. Caution should therefore be exercised when extrapolating these results to *in vivo* experiments. In general, the results from all of these preparations were in agreement, suggesting that the second-order maximum entropy model can consistently describe the origin of correlated states in a broad class of neural networks.

#### Comparison with previous work

Using the maximum entropy approach, Shlens et al. (2006) obtained substantially higher fractions of multi-information in retinal tissue ( $\sim 98\%$ ) than we report here for cortical tissue ( $\sim 88\%$ ). One reason for this difference could be that Shlens et al. (2006) only considered ensembles of parasol cells, whereas we did not restrict our ensembles to be composed of only one cell type. Circuit differences could also account for this disparity; the retina is primarily a sensory structure with only moderate feedback connections (Dowling, 1970), whereas the cortex is primarily an association structure with dense feedback connections (Braitenberg and Schuz, 1998). Our results are more similar to those of Schneidman et al. (2006) in terms of the fraction of multi-information and the general trends in local fields. Nevertheless, these findings demonstrate that pairwise interactions determine the preponderance of spatial correlation structure at the network level in cortical tissue.

Despite these general similarities, there were some notable differences between the present findings and those reported previously. The interaction strengths,  $J$ , tended toward zero with increasing ensemble size in cortical data. In retinal data,  $J$  showed no significant decline (Schneidman et al., 2006). Although decreases in  $J$  would be expected to lead to less constrained ensembles, the average clique size was smaller for cortical data than for retinal data. If taken at face value, this result would suggest that

cortical ensembles are less flexible than retinal ensembles. However, these disparities could reflect the effects of being driven by an external stimulus or differences in circuitry. Activity from cortical networks presented here was exclusively generated by spontaneous activity, whereas the retinal networks in the study by Schneidman et al. (2006) were shown natural movies. Moreover, the organotypic cultures used here contained portions of cortex, striatum and thalamus, allowing for the formation of long-range synaptic loops between brain structures. Future work should provide more information about the effects of spontaneous activity and circuit structure on the performance of the maximum entropy model.

Another interesting finding is that the model accounted for higher fractions of multi-information in LFP data than in spike data. Although both Schneidman et al. (2006) as well as Shlens et al. (2006) originally developed their models to describe correlated states among spikes, there is no reason why their approach would not be more generally applicable. The LFP signals that were used in the present work showed sharp negative peaks and therefore could be considered population spikes, indicative of groups of neurons firing action potentials synchronously near an electrode. This would be consistent with the finding that the clique sizes obtained for LFPs were significantly smaller than those obtained for spikes, and this was true for data at both bin widths. On average, spike cliques were approximately four times larger than LFP cliques, suggesting that multiple spikes contributed to each LFP. If this is the case, then the LFPs here could be thought of as a type of coarse-grained measurement of spiking activity.

### Implications for dynamic neural network models

The main novelty of this work is the finding that transitions between correlated states cannot be captured by merely concatenating states from the second-order maximum entropy model. Why might this be the case? Temporal correlations between states could be caused by delayed synaptic interactions between neurons. Because the present model did not incorporate information about delays, it may not have been equipped to predict transitions between correlated states.

This is an important issue that could be addressed by an improved version of the model. There have been numerous reports of temporal patterns of activity at the network level (Lindsey et al., 1997; Prut et al., 1998; Villa et al., 1999; Chang et al., 2000; Martignon et al., 2000; Beggs and Plenz, 2004; Ikegaya et al., 2004; Segev et al., 2004). Because our results also indicate that temporal correlations are a common feature of cortical networks, future models should be adapted to account for this. Our findings suggest that the present model captures spatial correlation structure and even sequence size distributions quite well (Fig. 8, right column). It could therefore serve as a foundation for a more complete model that would consider spatiotemporal relationships.

We examined the time-lagged correlation as a first step toward temporally extending the model. This variable indicated when the length distributions from the concatenation procedure differed significantly from the length distributions from the actual data. In other words, strong time-lagged correlations tended to indicate long temporal sequences. Because of this, the time-lagged correlation might help to define a transition energy between states in an extended model:

$$E(V^t \rightarrow W^{t+1}) = - \sum_{i=1}^N \sum_{j=1}^N K_{ij} \sigma_i^t \sigma_j^{t+1} \quad (25)$$

where  $V^t$  is a state at time  $t$ ,  $W^{t+1}$  is a state at time  $t + 1$ , and  $K_{ij}$  would be a temporal coupling term between neuron  $i$  at time  $t$  and neuron  $j$  at time  $t + 1$ . Such an equation would have to be appropriately constrained, but approaches like this could be a fruitful topic for future work (Dewar, 2003).

### References

- Amari S (2001) Information geometry on hierarchy of probability distributions. *IEEE Trans Inform Theory* 47:1701–1711.
- Baker RE, Corner MA, van Pelt J (2006) Spontaneous neuronal discharge patterns in developing organotypic mega-co-cultures of neonatal rat cerebral cortex. *Brain Res* 1101:29–35.
- Banker G, Goslin K (2002) *Culturing nerve cells*. Cambridge, MA: MIT.
- Bartho P, Hirase H, Monoconduit L, Zugaro M, Harris KD, Buzsaki G (2004) Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. *J Neurophys* 92:600–608.
- Beggs JM, Plenz D (2003) Neuronal avalanches in neocortical circuits. *J Neurosci* 23:11167–11177.
- Beggs JM, Plenz D (2004) Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *J Neurosci* 24:5216–5229.
- Bettencourt LM, Stephens GJ, Ham MI, Gross GW (2007) Functional structure of neuronal networks grown in vitro. *Phys Rev E Stat Nonlin Soft Matter Phys* 75:021915.
- Braitenberg V, Schüz A (1998) *Cortex: statistics and geometry of neuronal connectivity*. Heidelberg: Springer.
- Brewer G, Torricelli JR, Evege EK, Price PJ (1993) Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. *J Neurosci Res* 35:567–576.
- Chang EY, Morris KF, Shannon R, Lindsey BG (2000) Repeated sequences of interspike intervals in baroresponsive respiratory related neuronal assemblies of the cat brain stem. *J Neurophysiol* 84:1136–1148.
- Cossart R, Aronov D, Yuste R (2003) Attractor dynamics of network UP states in the neocortex. *Nature* 423:283–288.
- Cover TM, Thomas JA (1991) *Elements of information theory*. New York: Wiley.
- Darroch HN, Ratcliff D (1972) Generalized iterative scaling for log-linear models. *Ann Math Stat* 43:1470–1480.
- Dewar R (2003) Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *J Phys A Math Gen* 36:631–641.
- Dowling JE (1970) Organization of vertebrate retinas. *Invest Ophthalmol* 9:655–680.
- Etyan D, Marom S (2006) Dynamics and effective topology underlying synchronization in networks of cortical neurons. *J Neurosci* 26:8465–8476.
- Fischer Y, Wittner L, Freund TF, Gähwiler BH (2002) Simultaneous activation of gamma and theta network oscillations in rat hippocampal slice cultures. *J Physiol (Lond)* 539:857–868.
- Frechette ES, Sher A, Grivich MI, Petrusca D, Litke AM, Chichilnisky EJ (2005) Fidelity of the ensemble code for visual motion in primate retina. *J Neurophysiol* 94:119–135.
- Haldeman C, Beggs JM (2005) Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Phys Rev Lett* 94:058101.
- Ikegaya Y, Aaron G, Cossart R, Aronov D, Lampl I, Ferster D, Yuste R (2004) Synfire chains and cortical songs: temporal modules of cortical activity. *Science* 304:559–564.
- Jaynes ET (1957) Information theory and statistical mechanics. *Phys Rev* 106:62–79.
- Johnson NL, Kotz S, Balakrishnan N (1997) *Discrete multivariate distributions*. New York: Wiley.
- Landau LD, Lifshitz EM (1958) *Statistical physics*. London: Pergamon.
- Lindsey BG, Morris KF, Shannon R, Gerstein GL (1997) Repeated patterns of distributed synchrony in neuronal assemblies. *J Neurophysiol* 78:1714–1719.
- Litke AM, Bezawiff N, Chichilnisky EJ, Cunningham W, Dabrowski W, Grillo AA, Grivich M, Grybos P, Hottowy P, Kachiguine S, Kalmar RS, Mathieson K, Petrusca D, Rahman M, Sher A (2004) What does the eye tell the brain?: Development of a system for the large-scale recording of retinal output activity. *IEEE Trans Nucl Sci* 51:1434–1440.
- Martignon L, Deco G, Laskey K, Diamond M, Freiwald W, Vaadia E (2000)

- Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Comp* 12:2621–2653.
- Nirenberg S, Victor J (2007) Analyzing the activity of large populations of neurons: how tractable is the problem? *Curr Opin Neurobiol* 17:397–400.
- Petermann T, Lebedev M, Nicolelis M, Pleniz D (2006) Neuronal avalanches *in vivo*. *Soc Neurosci Abstr* 32:539.1.
- Potter SM, DeMarse TB (2001) A new approach to neural cell culture for long-term studies. *J Neurosci Methods* 110:17–24.
- Prut Y, Vaadia E, Bergman H, Haalman I, Slovin H, Abeles M (1998) Spatiotemporal structure of cortical activity: properties and behavioral relevance. *J Neurophysiol* 79:2857–2874.
- Ruelle D (2004) Conversations on nonequilibrium physics with an extraterrestrial. *Phys Today* 57:48–53.
- Sanchez-Vives MV, McCormick DA (2000) Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat Neurosci* 3:1027–1034.
- Sasaki T, Matsuki N, Ikegaya Y (2007) Metastability of active CA3 networks. *J Neurosci* 27:517–528.
- Schiff SJ, Jerger K, Duong DH, Chang T, Spano ML, Ditto WL (1994) Controlling chaos in the brain. *Nature* 370:615–620.
- Schneidman E, Still S, Berry MJ, Bialek W (2003) Network information and connected correlations. *Phys Rev Lett* 91:238701.
- Schneidman E, Berry MJ, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440:1007–1012.
- Segev R, Benveniste M, Hulata E, Cohen N, Palevski A, Kapon E, Shapira Y, Ben-Jacob E (2002) Long term behavior of lithographically prepared *in vitro* neuronal networks. *Phys Rev Lett* 88:118102.
- Segev R, Baruchi I, Hulata E, Ben-Jacob E (2004) Hidden neuronal correlations in cultured networks. *Phys Rev Lett* 92:118102.
- Shlens J, Field GD, Gauthier JL, Grivich MI, Petrusca D, Sher A, Litke A, Chichilnisky EJ (2006) The structure of multi-neuron firing patterns in primate retina. *J Neurosci* 26:8254–8266.
- Shu Y, Hasenstaub A, McCormick DA (2003) Turning on and off recurrent balanced cortical activity. *Nature* 423:288–293.
- Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB (2005) Highly non-random features of synaptic connectivity in local cortical circuits. *PLoS Biol* 3:e68.
- Stoppini L, Buchs PA, Müller D (1991) A simple method for organotypic cultures of nervous tissue. *J Neurosci Methods* 37:173–182.
- Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W (1998) Entropy and information in neural spike trains. *Phys Rev Lett* 80:197.
- Tang A, Jackson D, Hobbs J, Chen W, Smith J, Patel H, Beggs JM (2007) A second-order maximum entropy model predicts correlated network states, but not their evolution over time. Abstract II-62, Computational and Systems Neuroscience, Salt Lake City, February.
- Tateno T, Kawana A, Jimbo Y (2004) Analytical characterization of spontaneous firing in networks of developing rat cultured cortical neurons. *Phys Rev E Stat Nonlin Soft Matter Phys* 65:051924.
- Traub RD, Whittington MA, Colling SB, Buzsáki G, Jefferys JG (1996) Analysis of gamma rhythms in the rat hippocampus *in vitro* and *in vivo*. *J Physiol (Lond)* 493:471–484.
- Van Pelt H, Corner MA, Wolters PS, Rutten WL, Ramakers GJ (2004) Long term stability and developmental changes in spontaneous network burst firing patterns in dissociated rat cerebral cortex cell cultures on multielectrode arrays. *Neurosci Lett* 361:86–89.
- Villa AE, Tetko IV, Hyland B, Najem A (1999) Spatiotemporal activity patterns of rat cortical neurons predict responses in a conditioned task. *Proc Natl Acad Sci USA* 96:1106–1111.
- Wagenaar DA, Madhavan R, Pine J, Potter SM (2005) Controlling bursting with closed-loop multi-electrode stimulation. *J Neurosci* 25:680–688.
- Wirth C, Luscher HR (2004) Spatiotemporal evolution of excitation and inhibition in the rat barrel cortex investigated with multielectrode arrays. *J Neurophysiol* 91:1635–1647.
- Wu JY, Guan L, Tsau Y (1999) Propagating activation during oscillations and evoked responses in neocortical slices. *J Neurosci* 19:5005–5015.