Behavioral/Systems/Cognitive

Higher-Order Interactions Characterized in Cortical Activity

Shan Yu (余山),¹ Hongdian Yang (杨黉典),¹.² Hiroyuki Nakahara (中原裕之),³.⁴ Gustavo S. Santos,³ Danko Nikolić,⁵.⁶ and Dietmar Plenz¹

¹Section on Critical Brain Dynamics, Laboratory of Systems Neuroscience, National Institute of Mental Health, Bethesda, Maryland 20892-9663, ²Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland, 20742, ³Laboratory for Integrated Theoretical Neuroscience, RIKEN Brain Science Institute, Wako City, Saitama 351-0198 Japan, ⁴Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Midori-ku Yokohama, Kanagawa 226-8503 Japan, ⁵Department of Neurophysiology, Max Planck Institute for Brain Research, 60528 Frankfurt am Main, Germany, ⁶Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, 60325 Frankfurt am Main, Germany

In the cortex, the interactions among neurons give rise to transient coherent activity patterns that underlie perception, cognition, and action. Recently, it was actively debated whether the most basic interactions, i.e., the pairwise correlations between neurons or groups of neurons, suffice to explain those observed activity patterns. So far, the evidence reported is controversial. Importantly, the overall organization of neuronal interactions and the mechanisms underlying their generation, especially those of high-order interactions, have remained elusive. Here we show that higher-order interactions are required to properly account for cortical dynamics such as ongoing neuronal avalanches in the alert monkey and evoked visual responses in the anesthetized cat. A Gaussian interaction model that utilizes the observed pairwise correlations and event rates and that applies intrinsic thresholding identifies those higher-order interactions correctly, both in cortical local field potentials and spiking activities. This allows for accurate prediction of large neuronal population activities as required, e.g., in brain—machine interface paradigms. Our results demonstrate that higher-order interactions are inherent properties of cortical dynamics and suggest a simple solution to overcome the apparent formidable complexity previously thought to be intrinsic to those interactions.

Introduction

The mammalian cortex forms a complex network consisting of >10 10 neurons. Interactions between these neurons exist at many different scales ranging from microcircuits in cortical columns to cortical areas across the whole brain. Consequently, during perceptive, cognitive, and motor functions, cortical dynamics are characterized by spatially distributed coherent activity patterns that reflect these neuronal interactions (Gray et al., 1989; Abeles et al., 1993; Bressler et al., 1993; Vaadia et al., 1995; Riehle et al., 1997; Rodriguez et al., 1999). During the past two decades, the simplest form of interactions, i.e., pairwise correlations, has been well identified (Singer, 1999). However, the study of higherorder interactions, i.e., the ones that manifest only in triplets, quadruplets, etc., has been challenging. The fundamental problem is the number of potential higher-order interactions, which grows exponentially with system size, e.g., the number of neurons considered, and makes the interaction structure quickly intractable. So far it is not clear how to generally overcome this "curse of dimensionality." Consequently, the structure of higher-order interactions in cortical activity and its underlying mechanisms have remained largely unexplored.

Previous work suggested that pairwise interactions alone provide a rather complete picture of neuronal activity, potentially circumventing the necessity of identifying higher-order interactions. Schneidman et al. (2006) and Shlens et al. (2006) showed that pairwise interactions explain most of the spiking activity in small (≤10 neurons) retina networks, a finding that was soon extended to include larger systems (Cocco et al., 2009; Shlens et al., 2009; Ganmor et al., 2011a), in vivo cortical networks (Yu et al., 2008; Ohiorhenuan et al., 2010), and correlations among population activities measured in the local field potentials (LFPs) (Tang et al., 2008; Santos et al., 2010). In contrast, a number of more recent studies (Montani et al., 2009; Ohiorhenuan et al., 2010; Santos et al., 2010; Ohiorhenuan and Victor, 2011; Ganmor et al., 2011b) have identified significant higher-order interactions in neuronal spiking and LFP activities, e.g., third-order interactions among closely neighbored (<300 µm) cortical neurons (Ohiorhenuan et al., 2010; Ohiorhenuan and Victor, 2011) and interactions of up to the eighth order among retinal ganglion cells (Ganmor et al., 2011b). Clearly, the structure of higher-order interactions and their contributions to cortical dynamics are still open to debate. Here we show that pairwise interactions alone are not sufficient to identify neuronal avalanche dynamics (Beggs and Plenz, 2003), measured as ongoing cortical activity in awake monkeys based on the LFP. We then demonstrate that the incorporation of a specific structure of higher-order interactions that results from thresholding (Amari et al., 2003; Macke et al., 2011)

Received June 17, 2011: revised Oct. 13, 2011: accepted Oct. 18, 2011.

Author contributions: S.Y., H.Y., and D.P. designed research; S.Y. and D.N. performed research; H.N. and G.S.S. contributed unpublished reagents/analytic tools; S.Y. and H.Y. analyzed data; S.Y. and D.P. wrote the paper.

This study was supported by the Intramural Research Program of the National Institute of Mental Health. Experiments with cat recordings were supported by a Deutsche Forschungsgemeinschaft Grant NI 708/2-1, Hertie Stiftung, Max-Planck Society, and the Frankfurt Institute for Advanced Studies. We thank Andrew Mitz and Richard Saunders for help in data collection and Steven Wise for his support during the initial phase of the project.

Correspondence should be addressed to Dietmar Plenz, Section on Critical Brain Dynamics, Laboratory of Systems Neuroscience, National Institute of Mental Health, 6001 Executive Boulevard, Bethesda, MD 20892-9663. E-mail: plenzd@mail.nih.gov.

DOI:10.1523/JNEUROSCI.3127-11. 2011 Copyright © 2011 the authors 0270-6474/11/3117514-13\$15.00/0

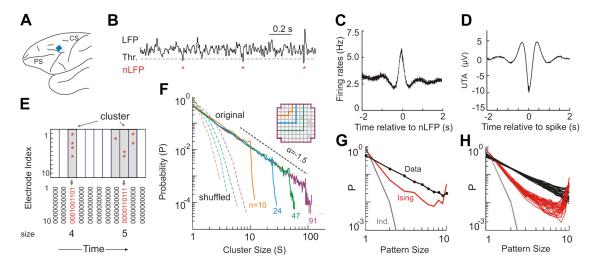


Figure 1. Power law organization of neuronal avalanches reflects higher-order interactions between locally synchronized neuronal groups. $\textbf{\textit{A}}$, Lateral view of the macaque brain showing the position of the multi-electrode array (blue square) in premotor cortex (not to scale). PS, principal sulcus; CS, central sulcus. $\textbf{\textit{B}}$, Example period of continuous LFP at a single electrode. Asterisks indicate peaks of negative deflections in the LFP (nLFP) that pass the threshold (Thr., broken line; -2.5 SD). $\textbf{\textit{C}}$, The nLFP indicates periods of increased local unit activity. Average PETH of unit firing relative to peak nLFP time obtained for LFP and unit activity recorded at the same electrode (ongoing activity; bin width, 2 ms; n=53 putative single units). $\textbf{\textit{D}}$, The average LFP waveform turns negative around the time of unit firing. Unit-triggered average (UTA) of the LFP waveform (n=53 units). Artifacts from spike waveform residuals due to incomplete LFP filtering were removed. $\textbf{\textit{E}}$, Identification of spatiotemporal nLFP clusters and corresponding spatial patterns. nLFPs that occur in the same time bin or consecutive bins of length Δt define a spatiotemporal cluster; the size is given by its number of nLFPs (top; gray area, two clusters of size 4 and 5). Patterns represent the spatial information of clusters only (bottom). $\textbf{\textit{F}}$, Neuronal avalanche dynamics are identified when the sizes of all clusters distribute according to a power law with a slope close of -1.5. Four distributions from the same original dataset (solid line) using different areas (inset), i.e., number of electrodes ($\textbf{\textit{n}}$), are superimposed. The power law distributions vanish for shuffled data (broken lines). A theoretical power law with a slope of -1.5 is provided as guidance to the eye (gray broken line). $\textbf{\textit{G}}$, The Ising model (red) fails to reconstruct the power law distribution of the 10-eletrode avalanche patterns obtained from the area indicated as orange in $\textbf{\textit{F}}$. For comparison, the pred

improves the accuracy of reconstructing neuronal avalanche dynamics by up to two orders of magnitude. The same method is then shown to significantly improve the approximation of ongoing spiking activity in awake monkeys as well as visually evoked spiking responses in anesthetized cats. Our results characterize higher-order interactions at different scales of cortical dynamics with an unprecedented efficiency and robustness and suggest a possible mechanism underlying their generation.

Materials and Methods

Electrophysiological recordings in monkeys. All experiments were performed in accordance with NIH guidelines for animal use and care. Ongoing LFP and spiking activity were recorded from two adult monkeys (Macaca mulatta; monkey A, female; monkey B, male). Multielectrode arrays (MEA; 96 channels; 10 \times 10 without corners, 400 μm interelectrode distance, 1 mm electrode length; Blackrock Microsystems) were chronically implanted in the arm representative region of the left premotor cortex (Fig. 1A). Approximately 30 min of ongoing LFP (1-100 Hz) and extracellularly recorded spiking activity (300-3000 Hz) were simultaneously obtained from each electrode while the animals were sitting alert in a primate chair with their heads fixed, but not engaged in a behavioral task. Spike sorting was performed off-line (Offline Sorter, Plexon). Putative single units were identified when (1) clearly isolated clusters were present in the PCA or waveform feature space and (2) the percentage of interspike intervals <1 ms was <0.2%. Based on these criteria, 56 (of 91 channels; monkey A) and 42 (of 78 channels; monkey B) putative single units were identified. Detailed results are shown for monkey A only. Qualitatively similar results for monkey B are summarized in Figures 7 and 10.

Electrophysiological recordings in cats. To study evoked spiking responses in vivo, we used datasets of visually evoked activities recorded from cortical area 17 in two, adult anesthetized cats published previously (Yu et al., 2008). In short, the animals were artificially ventilated and anesthesia was maintained with a mixture of 70% N₂O and 30% O₂, supplemented with 0.5–0.6% halothane. Extracellular spiking activity

was recorded by one or two silicon-based MEAs (4 \times 4 electrode array, 200 μ m interelectrode distance; Neural Nexus). The probes were inserted approximately perpendicular to the cortical surface to cortical depths of \sim 1 mm. Visual stimuli (presented by ActiveStim) consisted of full-contrast, drifting sinusoidal gratings that spatially covered the receptive fields of all recorded neurons. Each trial was completed with 3–4 s presentations of a drifting grating with a direction randomly chosen from a set of 12 directions (0 to 360° range; steps of 30 degrees). Approximately 30 min of visually evoked responses were recorded for each cat. Putative single units were identified off-line using a customized, PCA-based program. Three datasets were analyzed (two probes for cat A and one probe for cat B). Detailed results are only shown for cat A, probe one. Qualitatively similar results for the other two datasets are summarized in Figure 10.

LFP and avalanche analysis. Negative deflections in the LFP (nLFPs) were detected by applying a threshold at -2.5 SD of the LFP fluctuations estimated for each electrode separately (Fig. 1B). The nLFP peak times were then binned using a time window, Δt , to generate corresponding binary time series (1 for nLFP and 0 for the lack of it). Results shown are based on $\Delta t = 2$ ms. Similar results were obtained with other values of Δt ranging from 4 to 64 ms (data not shown). Spatiotemporal clusters of nLFPs, i.e., avalanches, were defined by consecutive bins that contained at least one nLFP at any of the recording sites (Beggs and Plenz, 2003). The size of a cluster was defined as the number of nLFPs in the cluster (Fig. 1 E). Cluster size distributions were obtained using linear bins (i.e., size equals 1, 2, 3, ...) and plotted in double-logarithmic coordinates for visual inspection. Shuffled time series of nLFPs were obtained by randomly switching interevent times between nLFPs for each site, i.e., electrode. This approach preserves the local average rates and interevent interval distributions of nLFPs, but removes spatiotemporal correlations among sites on the array.

Avalanche patterns and instantaneous nLFP patterns. The analysis of avalanches was simplified by collapsing all time bins within an avalanche to obtain a corresponding spatial pattern $\mathbf{x} = (x_1, x_2, x_3, \dots x_n)$, where n is the number of recording sites included in the analysis and $x_i = 1$ if at

least one nLFP occurred at site i and $x_i = 0$ otherwise (Fig. 1 E). These patterns will be referred to as "avalanche patterns" hereafter. Collapsed bins were replaced with the all zeros pattern such that the total duration of the dataset was preserved. We note that our results hold even if this replacement strategy is not used, as only <5% of bins were lost without replacement and our results do not take into account interpattern intervals (see also the analysis on instantaneous nLFP patterns below).

We also studied single bin nLFP patterns, obtained after binning nLFP peak times at temporal resolution Δt , but without concatenating successive bins with active sites into an avalanche. That is, $\mathbf{x} = (x_1, x_2, x_3, \dots x_n)$, where $x_i = 1$ if at least one nLFP occurred at site i and $x_i = 0$ otherwise. These spatial patterns will be referred as "instantaneous nLFP patterns" hereafter.

Because of the computational demands of the Ising model (see main text), most of our analysis was restricted to patterns obtained from a subset of n=10 electrodes if not stated otherwise. For spatially compact subsets (i.e., the neighboring cluster of electrodes), detailed results are shown. Qualitatively similar results for randomly chosen subsets are summarized in Figure 7.

Coherence potential analysis. The method was adopted from Thiagarajan et al. (2010) with slight modifications. Coherence potentials were identified if nLFP waveforms (baseline-to-baseline excursion) recorded si-

multaneously from different cortical sites were highly similar. For individual sites, nLFP waveforms were categorized according to their absolute peak amplitudes, a, binned at a resolution of 0.5 SD, i.e., a_0 = [0, ..., 0.5) SD, $a_1 = [0.5, ..., 1)$ SD, ..., $a_9 = [4.5, ..., 5)$ SD. For every amplitude level a_i, up to 300 randomly chosen nLFP waveforms were compared with simultaneously recorded waveforms from all other sites. The pairwise Pearson correlation coefficient, r, between nLFPs was computed to quantify waveform similarity. The fraction of sites, $F_{R\min}$, which had waveform similarities $r > R_{\min}$ was plotted against a_i . The resulting function represents a density function of coherence potentials, i.e., the extent of spatial coupling as a function of nLFP peak amplitude. This approach differs slightly from Thiagarajan et al. (2010), where a minimal amplitude threshold was used to categorize nLFPs, i.e., >0.5, >1, ..., >4.5 SD and the resulting sigmoid function represented cumulative probabilities of coherence potentials. We systematically varied R_{\min} from 0.1 to 0.9 to investigate the effects of its value on the results regarding spatial coupling and nLFP amplitude. To estimate the a priori probability of finding LFP waveforms with similarity larger than $R_{\rm min}$, for every <code>nLFP</code> analyzed, we randomly chose a length-matched LFP segment from the same channel and applied the identical analysis. The excess occurrence of nLFPs with high waveform similarity was quantified as the difference between the spatial coupling among nLFP waveforms and that of the length-matched, random LFP segments.

Spike analysis. To obtain activity patterns for spikes, spike trains were binned at temporal resolution Δt . "Instantaneous spike patterns" were taken as the spatial distribution of spikes within Δt without further temporal concatenation, as for the "instantaneous nLFP patterns." That is, $\mathbf{x} = (x_1, x_2, x_3, ..., x_n)$, where $x_i = 1$ if at least one spike generated by unit i and $x_i = 0$ otherwise. We used a $\Delta t = 20$ ms for ongoing activity in the monkey and a smaller $\Delta t = 4$ ms for evoked activities in the cat, because of the relatively high firing rates during evoked responses compared with ongoing activity (Yu et al., 2008). The range of Δt used here is consistent with previous studies using the Ising model (Schneidman et al., 2006; Yu et al., 2008; Ohiorhenuan et al., 2010), and similar results were obtained

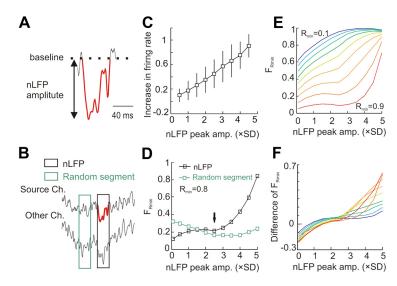


Figure 2. Coherence potential analysis suggests an intrinsic threshold in nLFP pattern formation. **A**, Coherence potential analysis utilizes the full baseline-to-baseline excursion of the nLFP waveform (red). **B**, The nLFP waveform at a source electrode is compared with the simultaneously recorded LFP at other electrodes (black rectangle). Similarities for each pairwise comparison are quantified by the Pearson correlation coefficient, r. Comparisons for randomly chosen, length-matched LFP segments from the source electrode serve as control (green rectangle). **C**, nLFP amplitude correlates with transient increase of local unit firing. Normalized increase of firing (see Materials and Methods) plotted as a function of nLFP amplitude (mean \pm SD; n=56 units). For all measured nLFP amplitudes, the increase of firing is significant ($p \ll 10^{-6}$; t test). **D**, Similarity in nLFP waveforms at distant sites increases nonlinearly with nLFP peak amplitude. Plotted are the fractions of electrodes on the array with high similarity r > 0.8, i.e., $R_{\min} = 0.8$, with the source electrode nLFP as a function of nLFP amplitude. Average increase in nLFP waveform similarity over all source channels (black; n=91). Expectation in similarity for random comparisons (green). Threshold (-2.5 SD; arrow) used for nLFP detection in the current neuronal avalanche analysis. **E**, The nonlinear coherence potential function is revealed for high-similarity requirements, i.e., $R_{\min} > 0.6$. Coherence potential probability plotted as a function of minimal similarity $r > R_{\min}$ (color coded). **F**, Difference between functions in **E** and expectation from random controls.

by varying Δt by at least two folds, i.e., 10 ms for ongoing activity and 8 ms for evoked activity (data not shown).

To study model performances as a function of the average level of pairwise correlation, subsets of units with strong pairwise correlations were identified (Figs. $10\,B$ and 11). For three-unit groups, we calculated the average pairwise correlation for all possible three-unit combinations and chose those 100-250 triplets with the highest average correlation. For 10-unit groups, we calculated the average pairwise correlation for all possible 10-unit combinations or up to 10^6 randomly chosen ones. Among those, we analyzed 15-30 groups with the highest average correlation value. We note that the majority of units (>60%) were included in the analysis of strongly coupled groups.

Peri-event time histogram of spikes. For each electrode, nLFPs within a given amplitude range $a_0 = [0.1, ..., 0.5)$ SD, $a_1 = [0.5, ..., 1)$ SD, ..., $a_9 = [4.5, ..., 5)$ SD were selected and used as triggers to construct peri-event time histograms (PETHs) of units recorded at the same electrode. Histograms were binned at 2 or 32 ms, converted to rates, and ranged from -2 to 2 s. The difference in PETH peak value and the unit's mean firing rate was obtained for each nLFP amplitude range, normalized by the maximal difference and averaged over all units (Fig. 2*C*). A bin width of 32 ms was used for the analysis shown in Fig. 2*C*, which reduces the effect of systematically overestimating peak PETH values due to "leaked" spike waveforms in the LFP.

Ising model. The Ising model (Schneidman et al., 2006) is expressed as $P(\sigma) = Z^{-1} \exp{(\Sigma_i h_i \sigma_i + \Sigma_{ij} J_{ij} \sigma_i \sigma_j)}$, where $P(\sigma)$ is the probability of the pattern $\sigma = (\sigma_1, \sigma_2, ..., \sigma_n)$. σ_i equals 1 or -1, representing the state of the ith element (1 for active and -1 for inactive). The normalization factor Z, intrinsic property h_i , and interaction J_{ij} were determined according to the experimentally observed rate $\langle \sigma_i \rangle$ and pairwise correlation $\langle \sigma_i \sigma_j \rangle$. The Ising model is a maximum entropy model with given $\langle \sigma_i \rangle$ and $\langle \sigma_i \sigma_j \rangle$. For 3- and 10-element systems, the Ising model was numerically solved using custom-written MATLAB code incorporating the MATLAB optimization toolbox function fsolve. For systems with up to 10 elements, the Ising model can be solved exactly with an error in fitting the

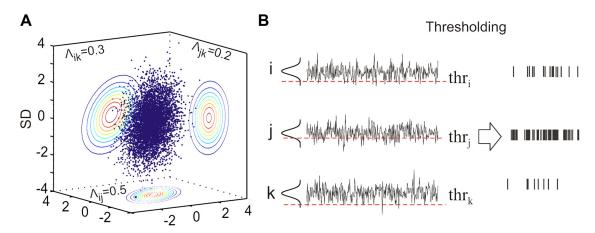


Figure 3. DG model. *A*, The distribution of a 3D Gaussian and its projections at individual 2D subspaces. Marginal distributions at a 2D space are represented by probability density contours. Λ, pairwise covariance. *B*, Converting continuous Gaussian variables to binary variables (events) by thresholding. Three Gaussian variables with pairwise covariance as specified in *A* and without higher-order interactions (left). Individual thresholds (thr_i, thr_j, and thr_k; broken lines) applied to each variable. The resulting trains of events, i.e., nLFPs or spikes, after thresholding (right). To fit the DG model to the data, the pairwise covariance (Λ) of the multidimensional continuous Gaussian and each threshold need to be adjusted such that the resulting binary variables have event rates and pairwise correlations identical to the data. We note that there is no direct correspondence between the threshold to extract nLFP patterns and the individual thresholds to adjust the rates of the binary variables.

rates and pairwise correlations smaller than 10 $^{-9}$. This results in almost identical lower-order statistics, i.e., event rates and pairwise correlations, between the Ising model and the data. For example, for the 10-electrode group shown in Figure 1G, the approximation error defined as $|v_{\rm Ising}-v_{\rm data}|/v_{\rm data}$, was $<10^{-10}$, where $v_{\rm data}$ is the value of the rates or pairwise correlations calculated from the data, and $v_{\rm Ising}$ is the corresponding value calculated from the pattern probabilities of the Ising model.

Dichotomized Gaussian model. The dichotomized Gaussian (DG) model (Amari et al., 2003; Macke et al., 2009, 2011) has a threshold operation based on multivariate Gaussian variables: $y_i = 1$ when $u_i > 0$ and $y_i = 0$ otherwise, where $\mathbf{u} = (u_1, u_2, ..., u_n) \sim N(\gamma, \Lambda)$ and γ is the mean and Λ is the covariance of the Gaussian variables, respectively. To match the rate, r, and covariance, Σ , of the observed binary variables, γ and Λ need to be adjusted according to $\gamma_{\rm i} = \Phi^{-1} \left(r_{\rm i} \right)$ and $\Lambda_{\rm ij}$ as the solution for $\Sigma_{ij} = \Phi_2(\gamma_i, \gamma_j, \Lambda_{ij}) - \Phi(\gamma_i) \Phi(\gamma_j)$, where Φ and Φ^{-1} are cumulative probability function of standard Gaussian distribution (Φ for 1D and Φ_2 for 2D) and its inverse function, respectively. Implementation of the model with MATLAB can be found in Macke et al. (2009). The pattern probabilities for the DG model were calculated using the cumulative distribution of multivariate Gaussians (MATLAB function mvncdf) or running simulations (for large systems). In terms of event rates and pairwise correlations, the fitting precision of the DG model was usually 10⁻³, which reaches or exceeds the criteria used in related studies (Ganmor et al., 2011a,b), but is inferior compared with that of the Ising model. This limit is mainly due to the noise introduced in estimating the cumulative probability distributions of the multivariate Gaussian, for which the tolerance error was set to $<10^{-4}$. Therefore, the superior performance of the DG model over the Ising model in approximating pattern probability and pattern size (see Results) cannot be attributed to a higher accuracy of the DG model in fitting event rates and pairwise

Model performance measures. To quantify the total interdependence of the data that can be explained when only considering pairwise interactions, we used an entropy-based approach (Schneidman et al., 2006). The entropy (H) was calculated as $H = \Sigma - P(\mathbf{x}) \log P(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, ..., x_n)$. $x_i = 1$ or 0/-1 indicates the status of element i to be active or inactive, respectively. Then, the proportion of correlation that can be explained by pairwise interactions is represented by the ratio $(H_{\mathrm{Ind}} - H_{\mathrm{Ising}})/(H_{\mathrm{Ind}} - H_{\mathrm{data}})$. Here, H_{Ind} is the entropy obtained from the "independent model" that matches event rates of the data, but assumes no correlation among elements, H_{Ising} is the entropy obtained from the Ising model and H_{data} is the entropy obtained from the original data.

The Jensen–Shannon (JS) divergence was used to quantify the performance of different models in approximating the data. The JS divergence

is a symmetric, bounded measure of the distance between two distributions and has been used in related studies (Schneidman et al., 2006; Ganmor et al., 2011a). More precisely, the JS divergence between probability distribution p and q is noted as $D_{\rm JS}$ ($p\|q$) and calculated as $D_{\rm JS}$ ($p\|q$) = 0.5 $D_{\rm KL}(p\|m)$ + 0.5 $D_{\rm KL}(q\|m)$, where m=(p+q)/2 and $D_{\rm KL}$ is the Kullback–Leibler divergence, calculated as $D_{\rm KL}(p\|q) = \sum p_i \log(p_i/q_i)$. A two-fold cross-validation (Santos et al., 2010) was used for model comparison. That is, time bins of the original dataset were randomly assigned to one of two sets. Model parameters were determined from one set only and predictions were made for the second set. In addition to the DG, Ising, and independent models, a "half-data" model was also included for comparison. The half-data model simply uses the results, i.e., pattern probabilities or pattern sizes, measured in one-half of the dataset to predict the other half.

In estimating the JS divergence, a decision had to be made regarding the handling of patterns that were not observed empirically. Simply assigning a probability of zero to nonobserved patterns introduces a systematic bias in assessing a model's performance, because this approach underestimates the true probability of these patterns. To avoid such bias, only the common patterns that were actually observed in both data and all models (including the half-data model) were taken into account in calculating $D_{\rm IS}$.

Three controls were conducted to ensure that our conclusions about model performances were not affected by the exclusion of nonobserved patterns. In the first control, we analyzed the DG model's prediction about nonobserved patterns, an aspect that is not reflected in $D_{\rm IS}$. We reasoned that a prediction for nonobserved patterns is consistent with the experimental finding, if the predicted pattern probability is considerably smaller than 1/N, where N is the total number of patterns observed. Indeed, the average prediction of the DG model for the probability of nonobserved patterns was $(0.46 \pm 0.04)/N$ (mean \pm SD; across 30 10-electrode groups analyzed in Fig. 4C). In the second control, we empirically tested how the number of nonobserved patterns affected our model comparisons. We systematically changed the sample size from 10% to 100% of the recording length of the original dataset and computed the ratio $D_{\rm JS}(p_{\rm data}\|p_{\rm Ising})/D_{\rm JS}(p_{\rm data}\|p_{\rm DG})$. We found that this ratio was consistently larger than 1 for small sample sizes (e.g., with 10% recording length, the ratio is ~4 for avalanche patterns and ~1.4 for instantaneous spike patterns) and increased further or remained relatively stable as the sample size increased. For example, for the full, 100% recording length, the ratio was >10 for avalanche patterns and \sim 1.6 for instantaneous spike patterns. This demonstrates the superiority of the DG model over the Ising model across varying proportions of nonobserved patterns. It further suggests that our findings are likely to hold even if sample size would be sufficiently increased to reliably estimate the probabilities of all potential patterns in a system. In our third and final control, we assigned zero probability to nonobserved patterns, which allowed us to include all patterns in the computation of $D_{\rm JS}$ for the DG and Ising models. Unlike the case for KL divergence, the additional zero-probability patterns increase the magnitude of the JS divergence, given that the corresponding probabilities are nonzero in the model. However, we found that this approach did not change our conclusions about model performances (data not shown).

The statistical significance of performance differences were assessed with the paired-sample Wilcoxon signed rank test.

Calculating interactions of different orders. Calculation of interactions in the three-element groups were based on full log-linear expansion (Amari, 2001; Nakahara and Amari, 2002): log $P(\mathbf{x}) = \sum \theta_i x_i + \sum \theta_{ij} x_i x_j + \sum \theta_{ijk} x_i x_j x_k - \psi$, where P(x) is the probability of the pattern $x = (x_1, x_2, \dots, x_n)$ x_3). x_i equals 1 or 0, representing the state of the *i*th element. ψ is a normalization factor. θ_i , θ_{ii} , θ_{ij} , and ψ were determined according to P(x), which is either experimentally observed (for the data) or directly calculated (for both the DG and Ising models). In principle, the calculations can be done for arbitrarily high orders, i.e., up to the system size. However, the amount of data needed for accurate calculation increases exponentially with higher order. Therefore, although interactions of higher orders larger than three exist in the DG model (Amari et al., 2003), and may be also in the data, we restricted our analysis to maximal third-order interactions given the limited length of our datasets.

Terminology. In the current study, "correlation" and "interaction" have different meanings and are usually not interchangeable. For random variables $x_1, x_2, ..., x_n$ correlation measures the degree of apparent dependency among them. For example, pairwise correlation refers to the quantity like $\langle x_i x_i \rangle$ (or $\langle x_i x_i \rangle - \langle x_i \rangle \langle x_i \rangle$). In contrast, interaction refers to the quantity θ as defined in the previous section, which does not have a simple relation to correlations. We also note that θ is derived from the observed statistics and does not necessarily reflect any physical, causal relation such as synaptic interaction between two or more elements. For the interested reader, correlation and interaction are related to the η and θ coordinates in information geometry, respectively (Nakahara and Amari, 2002).

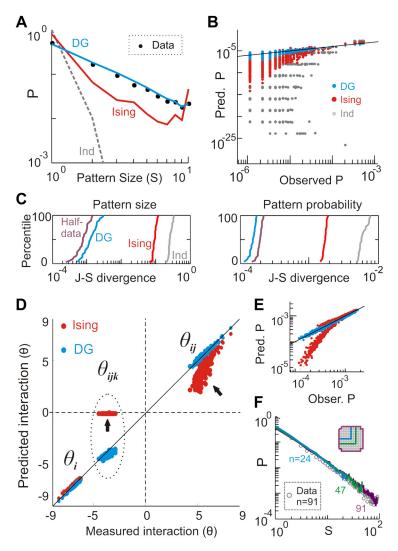


Figure 4. The DG model accurately reconstructs neuronal avalanche dynamics and predicts second- and third-order interactions in neuronal avalanches significantly better than the Ising model. A, The DG model (blue) reconstructs the power law in size distribution for avalanche patterns (dots) more accurately than the Ising model (red). Results for data, Ising model, and Independent (Ind) model replotted from Figure 16 for comparison. B, The DG model is also superior in predicting the probabilities of avalanche patterns (corresponding data from A). Observed pattern probability P is plotted against model predictions. Solid line indicates equality. The most common pattern (all zeros; inactive) is not shown for visual clarity. C, Quantification of model prediction demonstrates 1–2 orders of magnitude in improvement of the DG model over the Ising model. Cumulative distribution of JS divergence between the observed and predicted size distribution of avalanche patterns for n = 30 randomly chosen, spatially compact 10-electrode groups (left; monkey A; compare Fig. 1 H). Corresponding analysis for avalanche pattern probabilities (right). $One-half of randomly chosen bins was used to predict the other half of data (half-data; see Materials and Methods) \textbf{\textit{D}}, The DG model (half-data) and Methods) and Methods) \textbf{\textit{D}}, The DG model (half-data) and Methods) and Methods) and Methods) and Methods) and Methods) and Methods) and Methods (half-data) and Methods (half-data)$ accurately predicts second- (θ_{ii}) and third-order (θ_{iik}) interactions. In comparison, the Ising model is less accurate for second-order interactions (right arrow) and fails by design to predict third-order interactions (ellipsoid, arrow). Measured interactions of first to third order (θ_{ii} , θ_{iir} and θ_{iikr} respectively) are plotted against model predictions for n=98 randomly chosen 3-electrode groups (monkey A). E, Corresponding measured avalanche pattern probability plotted against model predictions (over all 3-electrode groups). Solid line indicates equality. F, The DG model accurately predicts the power law in the size distribution for avalanche patterns for systems much larger than 10 sites. Prediction for n = 24, 47, and 91 electrode sites are shown (compare Fig. 1 F).

Results

Higher-order interactions are essential for neuronal avalanche dynamics

We first examined the interaction structure in a well identified type of cortical dynamics—neuronal avalanches. Neuronal avalanches describe the spatiotemporal organization of synchronized activity in superficial layers of cortex. They have been demonstrated in spontaneous activity *in vitro* (Beggs and Plenz, 2003; Stewart and Plenz, 2006, 2008) and in ongoing activity *in vivo* (Gireesh and Plenz, 2008; Petermann et al., 2009; Hahn et al., 2010; Ribeiro et al., 2010). Importantly, avalanche size, *s*, distrib-

utes according to a power law, $P(s) \sim s^{\alpha}$ with exponent α close to -1.5, a hallmark of critical state dynamics (Plenz and Thiagarajan, 2007; Klaus et al., 2011). Both theoretical (Kinouchi and Copelli, 2006; Rämö et al., 2007; Tanaka et al., 2009) and empirical studies (Shew et al., 2009, 2011) suggest that avalanches optimize various aspects of information processing in cortical networks.

Using implanted 10×10 microelectrode arrays, we measured ~ 30 min of the ongoing LFP in premotor cortex of two alert macaque monkeys (Fig. 1A). Negative LFP deflections (Fig.

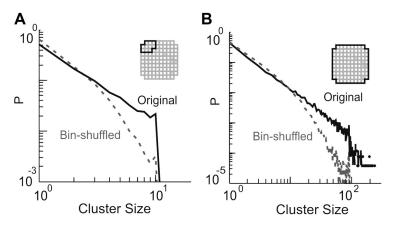


Figure 5. Demonstration of significant temporal correlations in avalanche dynamics. A–B, Size distribution for spatiotemporal clusters from bin-shuffled nLFPs (gray broken line) compared with the original avalanche size distribution (black solid line). For bin-shuffled data, temporal correlations beyond Δt , i.e., the duration of a single bin, are removed, which destroys the power law in cluster sizes. A, 10-electrode group. B, 91-electrode group. Insets, areas used for analysis.

 $1\,B$), which correlated in our recordings with neuronal firing (Fig. $1\,C$,D) (Petermann et al., 2009), were identified using an amplitude threshold a_{\min} of -2.5 SD of the LFP calculated for each electrode separately. Importantly, when nLFPs on the array were concatenated into spatiotemporal nLFP clusters (Fig. $1\,E$; top, temporal resolution $\Delta t = 2$ ms; see also Materials and Methods), the cluster size s distributed according to a power law with an exponent of -1.5 and the distribution showed finite-size scaling, i.e., the cutoff changed systematically with array size (Fig. $1\,F$) (Klaus et al., 2011). The power law and corresponding exponent identify the ongoing activity as neuronal avalanche dynamics. The power law indicates the presence of long-range correlations between cortical sites and, accordingly, is destroyed when nLFPs are shuffled randomly (Fig. $1\,F$, broken lines).

To investigate whether pairwise interactions are sufficient to explain the power law in avalanche sizes, we first constructed a maximum entropy model (Schneidman et al., 2006) based on the neural activity. Also known in physics as the "Ising" model, this is a widely used method that only utilizes the observed first order, i.e., event rate, and second order, i.e., pairwise correlation statistics. It ensures that no higher-order interactions are taken into consideration for reconstruction of the observed activities. The exact solution of the Ising model is computationally highly demanding for large neuronal groups (Cocco et al., 2009; Ganmor et al., 2011a,b), which limits our reconstruction of avalanche dynamics to arrays of size 10. The Ising model intrinsically lacks temporal structure and interactions are assumed to be instantaneous, whereas activities at different times are independent. To study neuronal avalanches with this model, we mapped the spatiotemporal organization of avalanches into corresponding spatial avalanche patterns (Fig. 1 E, bottom; see also Materials and Methods). This approach is justified because nLFPs at a single site rarely recur during an avalanche as indicated by the sharp cutoff in avalanche size distributions at system size in vitro as well as in vivo (Beggs and Plenz, 2003; Petermann et al., 2009; Klaus et al., 2011). Accordingly, size distributions of spatial avalanche patterns well matched the power law distribution of spatiotemporal avalanches (compare Figs. 1 F, G and 4F).

In line with previous studies (Tang et al., 2008; Santos et al., 2010), the Ising model captured >80% of interdependence for 10-electrode avalanche patterns (see Materials and Methods). More specifically, \sim 85% of interdependence was captured for the example shown in Figure 1G and on average 83 \pm 2% (mean \pm SD)

was captured for 30 randomly chosen, spatially compact 10-electrode patterns (Fig. 1 *H*). However, the Ising model failed to reconstruct the power law distribution of avalanche patterns. It significantly underestimated the probability of mediumsized patterns and overestimated the probabilities of small and large patterns (Fig. 1*G*,*H*). We conclude that neuronal avalanche dynamics contain significant higher-order interactions beyond that of pairwise interactions.

Coherence potentials identify an intrinsic threshold for nLFP pattern generation

To identify the higher-order interactions required for the reconstruction of neuronal avalanches, we analyzed in detail the emergence of nLFPs among spatially dis-

tributed channels using our previously introduced coherence potential analysis (Thiagarajan et al., 2010). Coherence potentials have been identified in ongoing neuronal avalanche activity as dynamics with an intrinsic threshold at the neuronal population level (Thiagarajan et al., 2010). A coherence potential consists of highly similar nLFP waveforms (baseline-to-baseline excursion) that occur almost simultaneously at different cortical sites (Fig. 2A,B). Importantly, the spatial extent of a coherence potential depends nonlinearly on the amplitude of the participating nLFPs. More specifically, while local unit firing increases linearly with nLFP amplitude (Fig. 2C), the spatial coupling, i.e., fraction of cortical sites, F, that engage in a coherence potential follows a nonlinear function. Figure 2D demonstrates this function in our recordings for ongoing avalanche activity in monkey A using a minimal similarity requirement $R_{\min} = 0.8$. Its steepest change occurs for nLFP amplitudes ≈ -2.5 SD (Fig. 2*D*, arrow). A systematic exploration of R_{\min} reveals that the nonlinearity in F is robustly found for $R_{\min} > 0.6$ (Fig. 2*E*). The inflection point of the function is also clearly visible when correcting for the a priori random expectation of coherence potentials (Fig. 2F; see Materials and Methods). Accordingly, at an amplitude threshold of -2.5 SD, the majority of nLFPs in our avalanche analysis in Figure 1 were identified as coherence potentials (86 \pm 8%, R_{\min} = 0.8; average across all electrodes), i.e., there was at least one other site with a highly similar nLFP waveform. These results suggest a threshold at ~ -2.5 SD intrinsic to the generation of the spatiotemporal avalanche patterns shown in Figure 1, above which the similarity in LFP waveforms rapidly increases.

Accurate approximation of avalanche patterns by a simple parametric model that utilizes a threshold function

Theoretical studies proposed nonlinear, e.g., threshold, operations as a potential origin for higher-order interactions (Amari et al., 2003; Macke et al., 2011). Therefore, to investigate whether higher-order interactions in neuronal avalanches can be explained by thresholding, we fit our data with the DG model (Amari et al., 2003; Macke et al., 2009, 2011). Like the Ising model, the DG model is a stochastic model without free parameters, i.e., all parameters are determined by the observed nLFP rates at each electrode and pairwise correlations between nLFPs at different sites. This model is based on multidimensional Gaussian variables followed by a threshold operation that results in stochastic output sequences of binary events (Fig. 3). The DG

model is fit to the data by (1) adjusting the threshold level for each dimension to match the observed nLFP or spike rates and (2) adjusting the covariance for individual pairs of dimensions to ensure that the resulting binary time series have the same pairwise correlations as the original nLFP or spike sequences (Fig. 3) (see also Material and Methods and Macke et al., 2009 for details). Whereas the Ising model contains only pairwise interactions, the DG model introduces, in addition, higher-order interactions among the output binary variables due to the threshold operations.

We found that the power law in the size distribution of avalanche patterns was correctly reconstructed when using the DG model (Fig. 4A). More precisely, the JS divergence (see Materials and Methods) between the predicted distribution and the true distribution was reduced by \sim 98% (0.002 for the DG model vs 0.085 for the Ising model). Moreover, the DG model predicted the probability of avalanche patterns much more accurately than the Ising model (10 sites; Fig. 4B; 1.3×10^{-4} JS divergence for the DG model vs 1.8×10^{-3} for the Ising model). When applying the two models to 30 randomly chosen, spatially compact 10electrode groups, we found that the DG model consistently outperformed the Ising model (on average 68 × better for size distribution and 11 × better for pattern probability; Fig. 4C). With respect to pattern probability, the accuracy of the DG model was even significantly better than the prediction based on the data directly (Fig. 4C, right; see Materials and Methods). This is possible because the DG model requires relatively few parameters to be determined, thereby reducing estimation errors that originate from finite sampling (see below for more analysis regarding this property). Similar results were found for monkey B (Fig. 7*A*).

For a more detailed comparison between the two models, we directly calculated first-, second-, and third-order interactions in the original data and the corresponding predictions from the DG and Ising models. Indeed, the DG model accurately predicted these interactions for neuronal avalanches as shown in Figure 4D. Moreover, this analysis not only confirmed that the Ising model failed to predict any third-order interactions, as

expected (Fig. 4D, ellipsoid, arrow), it also demonstrated that the Ising model was much less accurate in identifying second-order interactions when compared with the DG model (Fig. 4D, upper right, arrow). As a result of the estimation error in the interaction terms, we found that the probabilities of avalanche patterns pre-

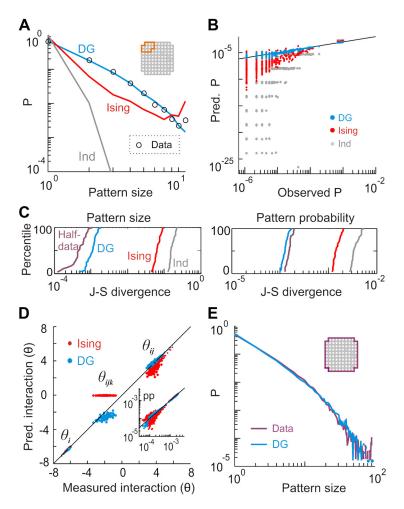


Figure 6. The DG model accurately predicts instantaneous nLFP patterns. The dataset used in A-E are the same as in Figure 4A-F, and all analysis was performed in a corresponding way, except that activity patterns were based on instantaneous nLFP patterns found within a single bin of duration Δt , without further temporal concatenation into spatiotemporal clusters. For details, see Figure 4 legend.

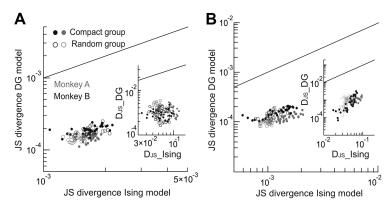
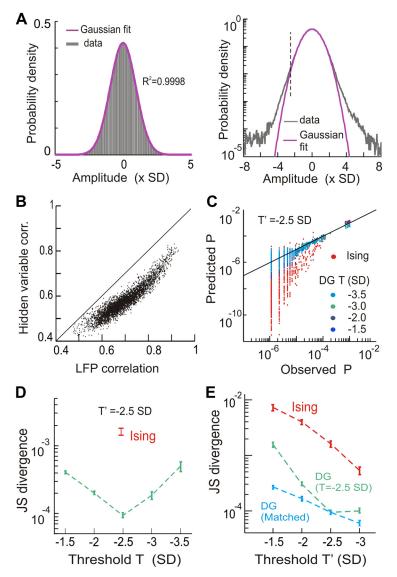


Figure 7. Summary of the comparison between the DG and Ising models in approximating nLFP patterns for both monkeys. The DG model provides much more accurate predictions for both avalanche patterns (**A**) and instantaneous nLFP patterns (**B**) (monkey A, gray; monkey B, black). JS divergence of Ising model in predicting pattern probabilities is plotted against that of the DG model (10-electrode groups; spatially compact: filled circle, 30 groups; random chosen: unfilled circle, 30 groups). Solid line indicates equality. Insets: Results for approximating the size distribution of avalanche patterns (**A**) and instantaneous nLFP patterns (**B**).

dicted by the Ising model significantly deviated from the observed probabilities, even in a small system with only three electrodes (Fig. 4E). Importantly, the DG model can be easily applied to larger systems (\gg 10 elements), for which the Ising model is computationally difficult. In Figure 4F, we demonstrate the significant of the significant of



strated accurate predictions of the DG model for neuronal avalanche dynamics, i.e., the pattern size distributions, which engage up to ~100 sites and cover ~16 mm² of cortex. We note that some of the results for large systems are well above the "perturbation regime" (Roudi et al., 2009), which is characterized as $Nr \ll 1$, where N is the number of elements and r is the average probability of event. Even for large Δt , when Nr > 1 (e.g., Nr = 1.8 and 3.5 for $\Delta t = 32$ and 64 ms, respectively), the DG model accurately captured the corresponding size distributions of the

91-electrode array (data not shown). Thus, our findings may generalize to even larger systems in which correlations become the dominant force in shaping the neuronal dynamics (Schneidman et al., 2006).

The contribution of temporal correlations to higher-order interactions in neuronal avalanches

To demonstrate the specific contribution of temporal correlations longer than Δt to higher-order interactions in avalanche dynamics, we used "bin-shuffling." More specifically, the instantaneous nLFP patterns (see Materials and Methods) at temporal resolution Δt were randomized with respect to bin position and then concatenated to obtain spatiotemporal clusters, i.e., bin-shuffled clusters. This approach maintains the original spatial correlation structure at temporal resolution Δt , but destroys temporal correlations longer than Δt . As shown in Figure 5, bin-shuffled spatiotemporal clusters do not show a power law in cluster sizes, demonstrating the presence of temporal correlations beyond Δt in neuronal avalanche dynamics. We then compared third-order interactions in the original avalanche patterns, $\theta_{ijk}^{\text{avl}}$, to those of bin-shuffled clusters, $\theta_{ijk}^{\text{binsh}}$. We expect that $|\theta_{ijk}^{\text{avl}}| > |\theta_{ijk}^{\text{binsh}}|$, which implies that a sizable amount of third-order interactions can be attributed to the temporal correlations intrinsic to neuronal avalanche dynamics. We further expect third-order interactions of instantaneous nLFP patterns, $\theta_{ijk}^{\text{inst}}$, to be in the range of $\theta_{ijk}^{\text{binsh}}$, which implies that the contribution to higher-order interactions from random temporal concatenation is small. Indeed, $|\theta_{iik}^{\text{avl}}| = 3.67 \pm 0.33$ (average and SD across 100, randomly chosen three-electrode groups from monkey A; the same as below), which was larger than $|\theta_{ijk}^{\text{binsh}}| = 2.07 \pm 0.46$ (by 78%; $p \ll 10^{-6}$; t test), while $|\theta_{iik}|^{\text{binsh}}$ was only slightly larger than $|\theta_{iik}|^{\text{inst}}$ $= 1.94 \pm 0.48$ (by 7%; p = 0.054; t test). These results demonstrate that higher-order interactions, exemplified here by thirdorder interactions, arise from spatial as well as temporal correlations in the original avalanche dynamics.

Accurate approximation of instantaneous LFP patterns by the DG model and possible mechanistic link between the model and the data

So far, we mapped temporal correlations beyond Δt into the spatial domain before applying the Ising or DG model. Here, we repeat our analysis for instantaneous nLFP patterns (see Materials and Methods), which do not involve temporal concatenation and which have been investigated in previous studies on neuronal

interactions (Tang et al., 2008; Santos et al., 2010). As shown in Figures 6 and 7*B*, the performances of the two models and the demonstration of higher-order interactions were highly similar to our results obtained with avalanche patterns.

Using the instantaneous nLFP patterns, we then investigated to which extent the DG model might reflect the actual data structure. Specifically, we examined whether the hidden layer of Gaussian variables in the model well approximates the recorded continuous LFP. First, we found that the 2D joint distribution of LFP amplitude from pairs of electrodes was well fit by a 2D Gaussian distribution ($R^2 = 0.9989 \pm 0.0003$ for 50 randomly chosen pairs). A similarly good fit was found for the marginal, i.e., 1D, distribution of LFP amplitude (Fig. 8A, left for an example; R^2 is 0.9997 \pm 0.0001 for all 91 electrodes). These results suggest that the multidimensional Gaussian used in the DG model provides a reasonable approximation for the actual LFP amplitude distribution. However, closer scrutiny revealed that this approximation only held for a certain range of LFP amplitudes. In fact, the LFP amplitude distribution systematically and significantly deviated from a Gaussian for amplitudes larger than $\pm 2-3$ SD. This deviation is not captured in the overall goodness of fit, but is evident in semilog plots (Fig. 8A, right) (Thiagarajan et al., 2010, their supplemental Fig. S10). We conclude that the Gaussian approximation of the DG model holds for LFP amplitudes up to $\sim \pm 2-3$ SD.

We then examined the relation between the hidden variables of the DG model and the correlation structure of LFPs using the full array. We note that when constructing the DG model, the covariances of the Gaussian variables are only determined by the pairwise correlations of the binary variables, i.e., nLFP patterns. In Figure 8 B, we showed that by fitting those correlations, the DG model also captured to a large degree the pairwise correlations among continuous LFP signals. More precisely, at a threshold of -2.5 SD as used in our analysis, the pairwise correlation in the LFP between two sites correlated well with that of the corresponding pair of Gaussian variables (Fig. 8 B; Spearman's r=0.92). On the other hand, the correlation between the hidden variables was systematically lower compared with the correlation of the corresponding LFP sites.

Next we asked whether the covariance matrix of the DG model estimated at nLFP threshold T can be used to predict nLFP patterns obtained with threshold T' (with proper rate adjustment). In Figure 8C and D, T' was fixed to -2.5 SD, while T was changed from -1.5 to -3.5 SD. In Figure 8 E, T was kept constant at -2.5SD, while T' was changed from -1.5 to -3 SD. Indeed, DG models constructed from mismatched, i.e., $T \neq T'$, covariance matrices provided fairly good fits (Fig. 8C) and yielded severalfold smaller JS divergence when compared with the Ising model (Fig. 8 D, E). On the other hand, the fitting accuracy was highest for the matched, i.e., T = T', DG model and deteriorated quickly when the distance between T and T' increased (Fig. 8 D, E). These results demonstrate that the covariance matrixes of the DG model are similar, but not identical across different nLFP thresholds. If mismatched and matched DG models had provided equally good fits, the DG model could have been considered a purely mechanistic model in which the hidden Gaussian variables represent the true LFP distribution and the nLFP patterns simply arise from thresholding. However, the threshold dependency suggests that this view is too simplistic. That is, while the analysis shows that the DG model can account for a large portion of the higher-order interactions, some systematic differences remain. This in turn suggests that a simple mechanism — local nonlinear transformation (e.g., thresholding) — plays a large role in the

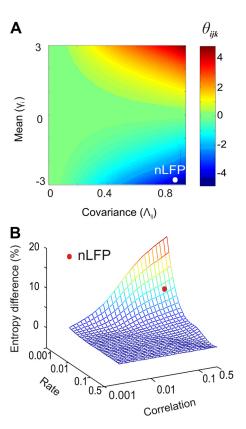


Figure 9. The magnitude of higher-order interactions introduced by thresholding depends on event rate and strength of pairwise correlations. $\textbf{\textit{A}}$, Third-order interactions in the binary variables cannot be neglected if, for the Gaussian variable, the covariance is strong and the mean is far below or above 0. Change in the magnitude of third-order interaction (θ_{ijk}) is shown as a function of the mean (γ) and the covariance (Λ) of an underlying 3D Gaussian, $\textbf{\textit{u}} \sim N\left(\gamma, \Lambda\right)$. θ_{ijk} was calculated for binary variables obtained by applying the threshold $u_i > 0$. The white dot marks the average mean and covariance of the hidden Gaussians estimated in the DG model for nLFPs (Fig. 4D). $\textbf{\textit{B}}$, Differences in pattern probabilities between the Ising and the DG models, here quantified as entropy difference, are most pronounced when, for the binary variables, the rate deviates far away from 0.5 (per bin) and pairwise correlations are strong, as is the case for avalanche patterns (high-rate regime produces a symmetric plot; data not shown). Entropy difference is plotted as a function of average rate and pairwise correlation. Dot (red) marks the average rate and pairwise correlation for avalanche patterns of monkey A (average of 91 channels and all pairs). For both $\textbf{\textit{A}}$ and $\textbf{\textit{B}}$, simulations assume homogenous rate and pairwise correlations for simplicity.

generation of higher-order interactions. However, the literal interpretation that a threshold is applied to an underlying multivariate Gaussian is likely an oversimplification for the generation of nLFP patterns (see Discussion for more details on this issue).

The nature of higher-order interactions generated by thresholding

To better understand the nature of the higher-order interactions that are introduced by thresholding in the DG model, we systematically varied the mean and covariance of a 3D Gaussian. The 3D Gaussian served as the layer of continuous variables in the model, before applying a fixed threshold (>0). We then investigated how these changes affected the third-order interactions in the resulting binary variables. As shown in Figure 9A, the thresholding procedure indeed introduced third-order interactions that were absent in the pre-thresholded Gaussian, especially when event probabilities were far above or below 0.5 and pairwise correlations were strong. We point out that an event probability r = 0.5 corresponds to a Gaussian mean $\gamma = 0$. We then varied the event probability and pairwise correlations for 10-element systems (bi-

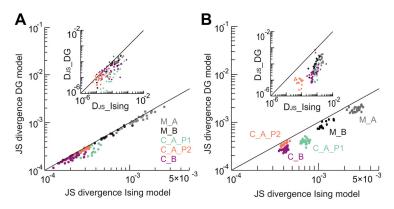


Figure 10. Summary of the performance of the DG and the Ising models in approximating the spiking activity for all five datasets. **A**, The DG model provides more accurate predictions than the Ising model for both pattern probabilities and pattern sizes, i.e., neuronal synchrony defined as the number of concurrent active units (inset). JS divergence of the Ising model in predicting pattern probabilities and number of concurrent spikes plotted against those of the DG model (15–30 randomly chosen, 10-unit groups for individual datasets; M_A, monkey A; M_B, monkey B; C_A_P1, cat A, probe one; C_A_P2, cat A, probe two; C_B, cat B). **B**, Similar to **A** but for neuronal groups with strong pairwise correlations (see Materials and Methods for details about selection). Fifteen to 30 10-unit groups were analyzed for individual datasets. For both **A** and **B**, the performance differences between the two models were statistically significant ($p \ll 10^{-6}$ for all comparisons; signed rank test; data combined across all datasets).

nary variables) and constructed both the DG and the Ising models using identical constraints. The difference between the two models indicates the total amount of higher-order interactions introduced by thresholding and was expressed as entropy difference (Fig. 9*B*). Consistent with our results shown in Figure 9*A* and previous findings (Macke et al., 2009, 2011), this difference was more pronounced when rates were low and pairwise correlations were strong, which is typically the case for nLFPs in ongoing activity (Petermann et al., 2009).

Accurate approximation of the ongoing and evoked spiking activity

We then extended our analysis to spiking activity of both ongoing and evoked cortical dynamics. We first analyzed extracellularly recorded spiking activity of up to 56 units during ongoing avalanche activity in the two awake monkeys. Similar to our nLFP analysis, in Figure 10A we show that the DG model significantly outperformed the Ising model in predicting both instantaneous spike patterns and their corresponding sizes, i.e., number of concurrent units, also referred to as neuronal synchrony (Montani et al., 2009). In line with the analysis shown in Figure 9, the improvement by the DG model over the Ising model was less pronounced compared with the LFP analysis, given the higher rates (0.075 vs 0.002, monkey A) and weaker pairwise correlations (0.04 vs 0.35) of spiking activity compared with nLFPs. Accordingly, when separately analyzing neuronal groups with relatively strong pairwise correlations (see Materials and Methods), predictions of the DG model for spike patterns, pattern sizes, and thirdorder interactions further improved over the Ising model (Figs. 10 B and 11A-C). We note that these groups included the majority of recorded units (>60%). Again, the DG model was readily extended to accurately predict the sizes of instantaneous spike patterns, i.e., spike synchrony, for large number of units (n = 56; Fig. 11 *D*). These results were not limited to ongoing activity only. We obtained similar results when analyzing spiking activity recorded from area 17 in anesthetized cats in response to grating stimuli (Figs. 10 and 11 E–H). The main difference between the two datasets referred to the range in pattern probabilities, which extended for the cat datasets down to 10⁻⁶ (compare Fig. 11A,E). We believe this difference is mainly due to the smaller bin width used in the analysis and high firing rate of evoked responses compared with the ongoing activity in the monkey. This analysis demonstrates the general applicability of the DG model to spontaneous and evoked spiking activities.

Efficient characterization of population activity

The approximation of higher-order interactions by a simple threshold operation allows for an efficient characterization of neuronal activity with relatively little data sampling. This is particularly desirable in studies of neural coding or brain-machine interfaces that require higher-order statistics of the population activity, such as pattern probabilities. While these statistics in principle can be directly obtained from the data (Fig. 4*B*), this approach becomes readily unfeasible for larger system sizes as outlined in the Introduction. The DG model only requires event rates and

pairwise correlations to approximate those pattern probabilities, which can be obtained within relatively short durations (<30 min in our analysis), and, importantly, those durations are independent of system size. The DG model is also simple enough to be treated analytically (Amari et al., 2003; Burak et al., 2009; Tchumatchenko et al., 2010; Macke et al., 2011), which allows for direct calculation of pattern probabilities. This advantage holds even for small systems (n = 10 elements) as shown in Figure 12 A. As typically done in brain-machine interface studies, we used a training dataset to make inferences about future data, i.e., the testing set. We randomly separated our 30 min recordings of nLFPs into two 15 min sets, which served as the testing and training set, respectively. We then systematically shortened the length of the training set and compared the performance in predicting the pattern probability for the testing set between the DG model and the direct sampling. The DG model increasingly outperformed direct sampling as the duration of the training sets shortened (Fig. 12A, C). To achieve the same prediction accuracy, the DG model required much shorter recording length (<1/3) compared with direct sampling (see Fig. 12B and corresponding legend for details). Similar advantages of using the DG model in characterizing pattern probabilities were also found for both ongoing and evoked spiking activities (Fig. 12C).

Discussion

Here we show that higher-order interactions are essential for understanding neuronal interdependence in the cortex, at the population level of neuronal avalanches and at the level of ongoing and evoked activities of individual neurons, especially when pairs of neurons are strongly coupled. This is consistent with recent findings that the inclusion of higher-order interactions yielded better approximation of LFP patterns (Santos et al., 2010) as well as neuronal synchrony (Montani et al., 2009). Our findings are also in line with the shortcomings reported for the Ising model to predict correlated activity of nearby neurons (Ohiorhenuan et al., 2010; Ohiorhenuan and Victor, 2011) or of large network activity (\sim 100 retina ganglion cells) in response to natural stimuli (Ganmor et al., 2011b). Moreover, our analysis provides new insights into these recent findings. For example, the relation between the pairwise and higher-order interactions as

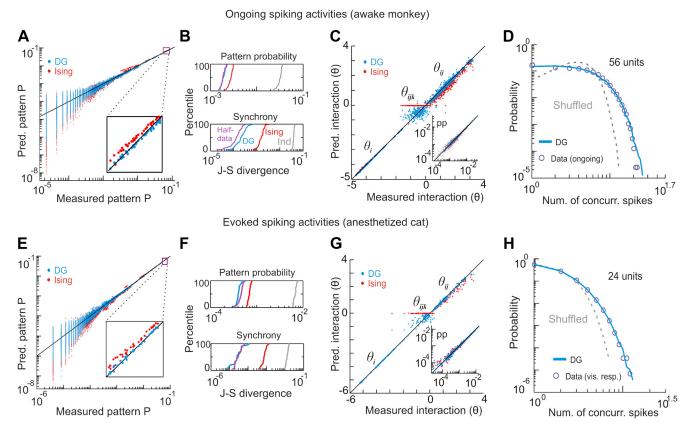


Figure 11. For strongly coupled subgroups, the DG model predicts spiking activities in ongoing and stimulus-evoked conditions significantly better than the Ising model. **A–C**, Ongoing spiking activity during avalanche dynamics in monkey A. **A**, Predictions in pattern probability by DG and Ising models (30 10-unit groups with strongest pairwise correlation). Zoomed-in view, showing that the DG model more accurately predicted high probability patterns (inset). **B**, Cumulative distributions of JS divergence for the groups shown in **A**. **C**, Measured and predicted interactions for >200 three-unit groups with strongest pairwise correlation. Measured (*x*-axis) and predicted (*y*-axis) pattern probabilities (pp) for the same three-unit groups (inset). For more details, see legend of Figure 4*B*–*D*. **D**, Distribution of concurrent spikes for all 56 neurons recorded in monkey A (circles; ongoing activity) and prediction made by the DG model (line). The distribution for the shuffled data was plotted for comparison (dotted line). The shuffling was made for individual neurons and the total number of spikes as well as the distribution of interspike intervals was preserved. **E–H**, Corresponding results for spiking activity evoked by drifting gratings recorded in area 17 of an anesthetized cat (cat A, probe 1).

revealed by the DG model (Fig. 9) provides a possible explanation for the phenomenon that the more pronounced third-order interactions are usually accompanied by strong pairwise interaction or correlation (Ohiorhenuan and Victor, 2011). In addition, the same mechanism may account for the specific failure of the Ising model to capture responses to natural stimuli (Ganmor et al., 2011b), as such responses are often characterized by sparse firing and strong pairwise correlations (Stüttgen and Schwarz, 2008; Jadhav et al., 2009). More generally, our results give a satisfactory account for the reported lack of higher-order interactions in many previous studies (Schneidman et al., 2006; Shlens et al., 2006; Tang et al., 2008; Yu et al., 2008), as those interactions are most pronounced in the regime of low rates and high pairwise correlation, which these studies often did not discriminate.

The significance of our study lies also in the clear demonstration that higher-order interactions in cortical dynamics are well approximated by a simple nonlinear operation, which has been suggested in a recent theoretical analysis (Macke et al., 2011). In the DG model, higher-order interactions are solely determined by the rates and pairwise correlations, so they do not serve as free parameters of the system and, therefore, they do not add extra complexity to the model. Importantly, the DG and Ising models have the same number of parameters. Thus, our results suggest a simple solution for taming the daunting complexity that was previously thought to be intrinsic to higher-order interactions.

Using only a quadratic number of parameters, the DG model provides a very compact description of data commonly expected to yield an exponential complexity of higher-order interactions. This indicates that the model captures, at least to some degree, the actual structure that generates the data (Baum, 2004). What would be the possible mechanistic link between the DG model and the neural process underlying the generation of spiking or nLFP patterns? For spiking activities, the correspondence is readily visible. The DG model assumes a Gaussian structure to which a threshold is applied. It approximates the membrane potential on which the cellular action potential generating mechanism operates. In fact, a Gaussian distribution of membrane potentials has been reported in intracellular recordings (Rudolph et al., 2004; Okun et al., 2010; Constantinople and Bruno, 2011), but see DeWeese and Zador (2006) for non-Gaussian distributions found under anesthesia. This finding is often used to support theoretical studies on neuronal processing that assume weak correlations between neurons: as a neuron usually receives a large number of synaptic inputs, central limit theorem indicates that the distribution of the membrane potential will be approximately Gaussian, provided that the correlation among those inputs is not too strong.

Our LFP and coherence potential analysis (Fig. 2; Thiagarajan et al., 2010) also suggest a potential mechanistic link between the DG model and nLFP patterns. LFP amplitudes are Gaussian distributed for small and median values, i.e., less than ± 2 –3 SD, but deviate from a Gaussian at larger values (Fig. 8*A*). The deviation corresponds with the nonlinear, threshold regime of extraordinary spatial coherence, the coherence potential (Fig. 2). These

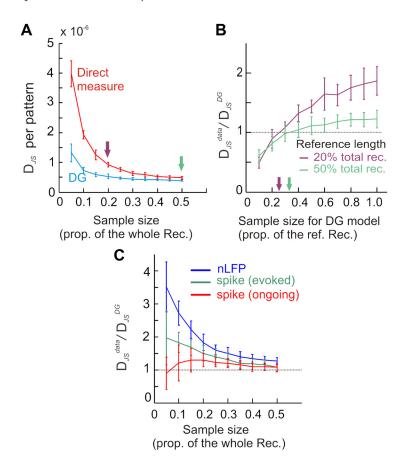


Figure 12. The DG model requires less data to characterize pattern probabilities compared with direct estimates. *A*, The DG model outperformed the direct sampling method in predicting pattern probabilities for nLFPs. Variably sized samples were drawn from 15 min of recording (training set) to predict pattern probability in another 15 min dataset (testing set). Performance quantified by average JS divergence per pattern in the testing set and plotted against the sample size taken from the training set. *B*, The DG model needs much less data to reach the same accuracy compared with direct sampling. Two reference recordings were chosen for the direct sampling method, 6 and 15 min in length, respectively (*A*, arrows). Then, various proportions of the reference recording were used for the DG model. The difference in performance of these two methods, measured by the ratio of JS divergence, was plotted against the amount of samples used by the DG model. The ratio of 1 (dotted line) indicates equal performance. Sample sizes to reach equal performance are marked by arrows. *C*, Sample size taken from the training set for both the nLFP and spiking activities plotted against the ratio of JS divergence. Almost all data points were significantly larger than 1 (signed rank test; p < 0.05), except for the first two points for ongoing spikes, which correspond to smallest sample sizes and, therefore, largest measurement errors. Total sample sizes were 9×10^5 , 9×10^4 , and 4.5×10^5 for nLFP, ongoing spikes, and evoked spikes, respectively. Data are represented as mean \pm SD. for all panels. Twenty randomly chosen 10-element groups for nLFP (monkey A), ongoing spikes (monkey A), and evoked spikes (cat A, probe 1) were analyzed.

coherence potentials have been identified as spatially distributed all-or-none population events of high fidelity in analogy to the action potential at the cellular level (Thiagarajan et al., 2010). As shown previously, the coherence potential threshold likely reflects intrinsic neuronal dynamics because it cannot be simply explained by (1) enhanced signal-to-noise ratio with high LFP amplitudes, (2) stronger volume conduction, or (3) common subcortical inputs as the dynamics also emerge in cortical cultures with neuronal avalanches (Thiagarajan et al., 2010). The neuronal mechanisms underlying such threshold-like behavior have yet to be identified. Because the nLFP amplitude correlates with the activity of local neuronal groups (Figs. 1C,D and 2C; Petermann et al., 2009), the putative threshold-like operation may reflect a nonlinear, recurrent recruitment of local suprathreshold activity that subsequently recruits distant sites (Douglas et al., 1995; Thiagarajan et al., 2010). On the other hand, despite such important mechanistic similarity between the DG model and the nLFP data, our exploration of the mismatched model suggested that the structural correspondence between the data and the model is not exact. This is not surprising as we know that the model is not intentionally constructed to reflect every process in the generation of the data. For example, in the DG model, all superthreshold values will result in the same binary state (active state; Fig. 3) while for nLFP patterns, only the peak of the negative LFP deflections were used to indicate the active state (i.e., the nonpeak values, although they might be superthreshold, will be indicated as inactive state). Furthermore, in the generation of avalanche patterns, there is a process mapping the spatiotemporal clusters to spatial patterns (Fig. 1), which is not used in the model.

Together, our results suggest that the DG model captured some essential features about the mechanisms underlying the formation of neuronal dynamics, at both the single unit and LFP level. Specifically, in terms of the origin of higherorder interactions, our results suggest that they are probably introduced by intrinsic thresholding such as in spike and nLFP pattern generation. However, to prove such a causal relation is awaiting future experiments, e.g., in which the level of pairwise correlation of the membrane potentials can be manipulated and the resulting changes in higher-order interactions among spike trains can be demonstrated to be consistent with the model's prediction.

For understanding spike-based neural code, we suggest studying the joint distributions of membrane potentials, as it may help to reveal fundamental constraints in neuronal processing. The multivariate Gaussian assumption is simple enough to allow a thorough analysis (Burak et al., 2009; Tchumatchenko et al., 2010) while we demonstrate that it accurately describes the data. If the actual distribution significantly deviates

from a multivariate Gaussian, it is likely that other forms of higher-order interactions will be present among the spike trains

The present results have direct implications, both conceptually and methodologically, for studying functional networks of neuronal systems (Bullmore and Sporns, 2009). On one hand, for binary time series such as spike trains, we have shown before that using interaction, rather than correlation, is necessary to extract the correct network architecture (Yu et al., 2008, their supplemental Fig. 2). Our current results demonstrate that the true pairwise interactions can only be captured by a model if potential higher-order interactions are taken into account. On the other hand, if one is interested in studying the interaction structure that underlies the observed binary time series (e.g., membrane potentials underlying spike trains) while this deeper level cannot be directly observed, the present results suggest that the DG model may serve as a useful tool to reveal this underlying structure. Studying this structure will significantly simplify the analysis of

neuronal functional networks as it dramatically reduces the number of interactions that need to be considered, which, to a first approximation, eliminates the need for considering any interactions above the second order. It allows the conventional graph theoretical analysis to be applied, which likely will yield more incisive results since the dimensionality of the problem has been much reduced.

References

- Abeles M, Bergman H, Margalit E, Vaadia E (1993) Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. J Neurophysiol 70:1629–1638.
- Amari S (2001) Information geometry on hierarchy of probability distributions. IEEE Trans Inform Theory 47:1701–1711.
- Amari S, Nakahara H, Wu S, Sakai Y (2003) Synchronous firing and higher-order interactions in neuron pool. Neural Comput 15:127–142.
- Baum EB (2004) What is thought?, Chap 6. Cambridge: MA MIT.
- Beggs JM, Plenz D (2003) Neuronal avalanches in neocortical circuits. J Neurosci 23:11167–11177.
- Bressler SL, Coppola R, Nakamura R (1993) Episodic multiregional cortical coherence at multiple frequencies during visual task-performance. Nature 366:153–156.
- Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186–198.
- Burak Y, Lewallen S, Sompolinsky H (2009) Stimulus-dependent correlations in threshold-crossing spiking neurons. Neural Comput 21:2269–2308.
- Cocco S, Leibler S, Monasson R (2009) Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. Proc Natl Acad Sci U S A 106:14058–14062.
- Constantinople CM, Bruno RM (2011) Effects and mechanisms of wakefulness on local cortical networks. Neuron 69:1061–1068.
- DeWeese MR, Zador AM (2006) Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. J Neurosci 26:12206–12218.
- Douglas RJ, Koch C, Mahowald M, Martin KA, Suarez HH (1995) Recurrent excitation in neocortical circuits. Science 269:981–985.
- Ganmor E, Segev R, Schneidman E (2011a) The architecture of functional interaction networks in the retina. J Neurosci 31:3044–3054.
- Ganmor E, Segev R, Schneidman E (2011b) Sparse low-order interaction network underlies a highly correlated and learnable neural population code. Proc Natl Acad Sci U S A 108:9679–9684.
- Gireesh ED, Plenz D (2008) Neuronal avalanches organize as nested thetaand beta/gamma-oscillations during development of cortical layer 2/3. Proc Natl Acad Sci U S A 105:7576–7581.
- Gray CM, König P, Engel AK, Singer W (1989) Oscillatory responses in cat visual-cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338:334–337.
- Hahn G, Petermann T, Havenith MN, Yu S, Singer W, Plenz D, Nikolić D (2010) Neuronal avalanches in spontaneous activity in vivo. J Neurophysiol 104:3312–3322.
- Jadhav SP, Wolfe J, Feldman DE (2009) Sparse temporal coding of elementary tactile features during active whisker sensation. Nat Neurosci 12:792–800.
- Kinouchi O, Copelli M (2006) Optimal dynamical range of excitable networks at criticality. Nat Physics 2:348–351.
- Klaus A, Yu S, Plenz D (2011) Statistical analyses support power law distributions found in neuronal avalanches. PLoS One 6:e19779.
- Macke JH, Berens P, Ecker AS, Tolias AS, Bethge M (2009) Generating spike trains with specified correlation coefficients. Neural Comput 21:397–423.
- Macke JH, Opper M, Bethge M (2011) Common input explains higherorder correlations and entropy in a simple model of neural population activity. Phys Rev Lett 106:208102.
- Montani F, Ince RAA, Senatore R, Arabzadeh E, Diamond ME, Panzeri S (2009) The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. Philos Transact A Math Phys Eng Sci 367:3297–3310.
- Nakahara H, Amari S (2002) Information-geometric measure for neural spikes. Neural Comput 14:2269–2316.
- Ohiorhenuan IE, Victor JD (2011) Information-geometric measure of 3-neuron firing patterns characterizes scale dependence in cortical networks. J Comput Neurosci 30:125–141.
- Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD

- (2010) Sparse coding and high-order correlations in fine-scale cortical networks. Nature 466:617–621.
- Okun M, Naim A, Lampl I (2010) The subthreshold relation between cortical local field potential and neuronal firing unveiled by intracellular recordings in awake rats. J Neurosci 30:4440–4448.
- Petermann T, Thiagarajan TC, Lebedev MA, Nicolelis MA, Chialvo DR, Plenz D (2009) Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. Proc Natl Acad Sci U S A 106:15921–15926.
- Plenz D, Thiagarajan TC (2007) The organizing principles of neuronal avalanches: cell assemblies in the cortex? Trends Neurosci 30:101–110.
- Rämö P, Kauffman S, Kesseli J, Yli-Harja O (2007) Measures for information propagation in Boolean networks. Physica D 227:100–104.
- Ribeiro TL, Copelli M, Caixeta F, Belchior H, Chialvo DR, Nicolelis MA, Ribeiro S (2010) Spike avalanches exhibit universal dynamics across the sleep-wake cycle. PLoS One 5:e14129.
- Riehle A, Grün S, Diesmann M, Aertsen A (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. Science 278:1950–1953.
- Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ (1999) Perception's shadow: long-distance synchronization of human brain activity. Nature 397:430–433.
- Roudi Y, Nirenberg S, Latham PE (2009) Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. PLoS Comput Biol 5: e1000380.
- Rudolph M, Piwkowska Z, Badoual M, Bal T, Destexhe A (2004) A method to estimate synaptic conductances from membrane potential fluctuations. J Neurophysiol 91:2884–2896.
- Santos GS, Gireesh ED, Plenz D, Nakahara H (2010) Hierarchical interaction structure of neural activities in cortical slice cultures. J Neurosci 30:8720–8733.
- Schneidman E, Berry MJ 2nd, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440:1007-1012.
- Shew WL, Yang H, Petermann T, Roy R, Plenz D (2009) Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. J Neurosci 29:15595–15600.
- Shew WL, Yang H, Yu S, Roy R, Plenz D (2011) Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. J Neurosci 31:55–63.
- Shlens J, Field GD, Gauthier JL, Grivich MI, Petrusca D, Sher A, Litke AM, Chichilnisky EJ (2006) The structure of multi-neuron firing patterns in primate retina. J Neurosci 26:8254–8266.
- Shlens J, Field GD, Gauthier JL, Greschner M, Sher A, Litke AM, Chichilnisky EJ (2009) The structure of large-scale synchronized firing in primate retina. J Neurosci 29:5022–5031.
- Singer W (1999) Neuronal synchrony: a versatile code for the definition of relations? Neuron 24:49–65; 111–25.
- Stewart CV, Plenz D (2006) Inverted-U profile of dopamine-NMDA-mediated spontaneous avalanche recurrence in superficial layers of rat prefrontal cortex. J Neurosci 26:8148–8159.
- Stewart CV, Plenz D (2008) Homeostasis of neuronal avalanches during postnatal cortex development in vitro. J Neurosci Methods 169:405–416.
- Stüttgen MC, Schwarz C (2008) Psychophysical and neurometric detection performance under stimulus uncertainty. Nat Neurosci 11:1091–1099.
- Tanaka T, Kaneko T, Aoyagi T (2009) Recurrent infomax generates cell assemblies, neuronal avalanches, and simple cell-like selectivity. Neural Comput 21:1038–1067.
- Tang A, Jackson D, Hobbs J, Chen W, Smith JL, Patel H, Prieto A, Petrusca D, Grivich MI, Sher A, Hottowy P, Dabrowski W, Litke AM, Beggs JM (2008) A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. J Neurosci 28:505–518.
- Tchumatchenko T, Geisel T, Volgushev M, Wolf F (2010) Signatures of synchrony in pairwise count correlations. Front Comput Neurosci 4:1.
- Thiagarajan TC, Lebedev MA, Nicolelis MA, Plenz D (2010) Coherence potentials: loss-less, all-or-none network events in the cortex. PLoS Biol 8:e1000278.
- Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A (1995) Dynamics of neuronal interactions in monkey cortex in relation to behavioral events. Nature 373:515–518.
- Yu S, Huang D, Singer W, Nikolić D (2008) A small world of neuronal synchrony. Cereb Cortex 18:2891–2901.