

ARTICLE

<https://doi.org/10.1038/s42005-019-0276-8>

OPEN

Fingerprints of a second order critical line in developing neural networks

Karlis Kandera^{1,2*}, Hyungsub Lee³, Nari Hong³, Yoonkey Nam³ & Ruedi Stoop^{1*}

Patterns of biological activity with properties similar to critical states of statistical mechanics have received much attention, as they were mostly seen as indicators of computational optimality. Commonly, a single regime around an isolated critical point is expected. Our experimental data and our network simulations of developing neural cultures indicate the possibility of transitions between different critical regimes. In the latter, the addition of further fundamental neurophysiological principles to the standard neurodynamics branching model generates steeper power laws that have been observed in various experiments. Our analysis exhibits two populations of neurons, each composed of inhibitory and excitatory sites, that have distinct dynamical and topological properties. This generates a line of second order critical points, similar to what is known from the thermodynamics of two-component alloys. An analysis of two major critical regimes found in the experiments suggests that different critical regimes may express distinct computational roles.

¹Institute of Neuroinformatics, ETH Zürich and University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland. ²Institute for Building Materials, ETH Zürich, Stefano-Francini-Platz 3, 8093 Zürich, Switzerland. ³Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea. *email: kkanders@ini.phys.ethz.ch; ruedi@ini.phys.ethz.ch

Collective dynamics of neurons¹ underlie cognitive functions and behavior of higher animals². Salient expressions of collective neural dynamics are continued spontaneous activities that modulate neural responses to external stimuli, where they often reflect expectations of future stimuli³. However, also spontaneous spatiotemporal patterns of neural activity, commonly termed “neural avalanches”, that have no characteristic scale (i.e., the distributions of avalanche sizes and lifetimes follow power laws), have been identified, e.g., in electrophysiological *in vitro* recordings^{4,5}, *in vivo* recordings^{6,7}, electroencephalogram recordings⁸, and in functional magnetic resonance imaging⁹.

Power laws are generated if a system is at a critical point of its system class. Critical states are typically associated with (continuous or second order—we use the terms here synonymously—) phase transitions that separate phases of qualitatively distinct dynamical behavior (e.g., ordered vs. disordered dynamics). Due to the diverging correlation lengths at critical states, systems near criticality show a high susceptibility to external stimuli. Seeing this as a favorable property of complex natural systems, it has been conjectured that neural systems operate preferentially in the close vicinity of a critical point^{10–14}. Why complex systems tend to be poised close to a critical point at the doorstep to instability if the system is slowly driven from outside, has been put forward by the hypothesis of self-organized criticality¹⁵. Power laws from critical systems are robust in the following sense: they are a property of the system class considered, and independent of the microscopic details of an individual system (i.e., they are “universal”). Systems from a neighborhood of the critical point deviate from critical systems again by simple class-specific scaling laws, expressing the distance of the system from the critical point. Moreover, the macroscopic behavior can accurately and efficiently be described by a very few relevant observables, which is beneficial for the understanding, modeling and prediction of systems¹⁶. Whereas the functional role of criticality has recently been pinned down for the peripheral auditory system¹⁷, the role of criticality in the brain is still actively debated^{18–20}. Presently, the general opinion states that at a critical state, the brain might optimize functional properties relevant for information processing (such as information transmission, information capacity, and response flexibility^{21,22}).

Direct demonstrations that real world power laws are the consequence of criticality are notoriously difficult (and rare), as the ideal matching would require to achieve the thermodynamic limit and observe a reliable divergence behavior at the critical point. In the experimental context, this is generally not possible. For the real world, the concept of criticality requires therefore adaptation, where the stability properties of critical regimes prove helpful to soften and generalize the “hard” criteria of criticality from statistical mechanics. As a result, a number of non-ideal biological real-world data and data from real-world models were successfully mapped to models of statistical mechanics and the corresponding universality classes could be identified. In previous experimental studies⁴, the scaling exponents associated with avalanche size and lifetime distributions, $p(S) \sim S^{-\tau}$ and $p(T) \sim T^{-\alpha}$, respectively, were mostly found to be $\tau \approx 1.5$ and $\alpha \approx 2.0$. These values match the theoretically expected critical exponents of the mean-field critical branching process^{23,24}, suggesting that biological networks may operate at a critical point characterized by a marginal propagation of activity, separating phases of quickly decaying and exploding runaway activity⁴. Later studies, however, provided evidence for a considerable variation of scaling exponents, e.g., across the interval [1.5, 2.6] for most prominent and easiest to measure exponent τ ^{4,10,25,26}. It has been argued that some of these deviations could be artefacts introduced by the experimental measurements (e.g., spatial undersampling of

neural activity^{13,27}) or by the computational avalanche extraction procedures^{4,13}.

One natural focus in the study of criticality in neuronal networks has been the development process of neuronal cultures, where in an earlier study²⁵ a critical regime with branching-type exponents was observed in the 8th week of development. In this study, the authors focused on the late development, while here we present results on the early development, i.e., after the first week *in vitro*. The distinct temporal foci affected several experimental parameters, such as the used culture neural densities (5000 cells mm^{-2} vs. our density of 800 cells mm^{-2}) and necessitated constraints in data acquisition. More recently, a study of dissociated hippocampal neurons co-cultured with glia cells (from newborn P0 Sprague Dawley rats) evidenced a second criticality regime with substantially larger exponents ($\tau \approx 2.2$, $\alpha \approx 3.3$)²⁸, in addition to the “expected” critical exponents. The observation of the second critical regime was, however, under condition of a specific pharmacological treatment (5M4Hfolate), which left open (in the work and to the reader) the question of how much importance should be attributed to this new regime.

Results

A model of network criticality with a line of critical regimes.

Maturation of *in vitro* neural networks is characterized by a gradual increase of network size and of the strength of the synaptic connections among the neurons. The effects of dynamic facilitation and depression when included in the classical branching model (see Methods section) are exhibited in Fig. 1. Early network maturation is represented in our modeling by an increase of the branching parameter (from $\sigma \approx 0.35$ to $\sigma \approx 0.9$), accompanied by a mild facilitation decrease (from $\Delta_\phi = 0.002$ to $\Delta_\phi = 0.0015$). The adjustment of facilitation is motivated by the observed decrease of neuronal excitability during maturation^{29,30}. In Fig. 1a), the effect of this ‘balanced’ facilitation is compared to the behavior of the original branching model (known to exhibit criticality at $\sigma = 1$). Moving from smaller to larger baseline branching parameters σ reduces the generated power-law exponents τ . For $\{\Delta_\phi, \sigma\} \rightarrow \{0, 1\}$, the value of τ settles to a value close to that of the classical branching model $\tau \sim 1.5$. In particular, the maturation paradigm (chosen to closely correspond to our experimental data shown later) shows avalanche size distribution transitions from (1) a subcritical phase (at $\sigma \approx 0.35$, with avalanches much smaller than system size), to (2) an early critical regime (at $\sigma \approx 0.61$, with a heavy-tailed distribution approximating a power law with $\tau \approx 2.2$). This regime then is followed (3) by a supercritical phase (at $\sigma \approx 0.71$, with a hump at the end of the distribution), giving (4) way to a “late” criticality regime (at $\sigma \approx 0.81$, with a shallower power law of $\tau \approx 1.65$), until a supercritical phase exhibiting a pronounced hump is reached. Through facilitation, the effective branching parameter $\bar{\sigma}_n$ (defined as the average sum of the outgoing excitation probabilities of each node at time step n) becomes a dynamic quantity exhibiting large fluctuations (Fig. 1b). A temporally variable branching parameter, based on continually altering neural activity embracing on top of background network activity the history of previous activity via short-term synaptic plasticity³¹, provides a more realistic description of biological neural networks³². The broad distributions of $\bar{\sigma}_n$ (Fig. 1c) obtained in this way seem to play an essential role in the generation of heavy-tailed avalanche size distributions. A change of σ entails substantial changes in the network’s configuration and topology (Fig. 1c–e). Strongly fluctuating network interactions also suggest that data-estimated branching parameters may not be too reliable for predicting a system’s distance to criticality^{4,14}.

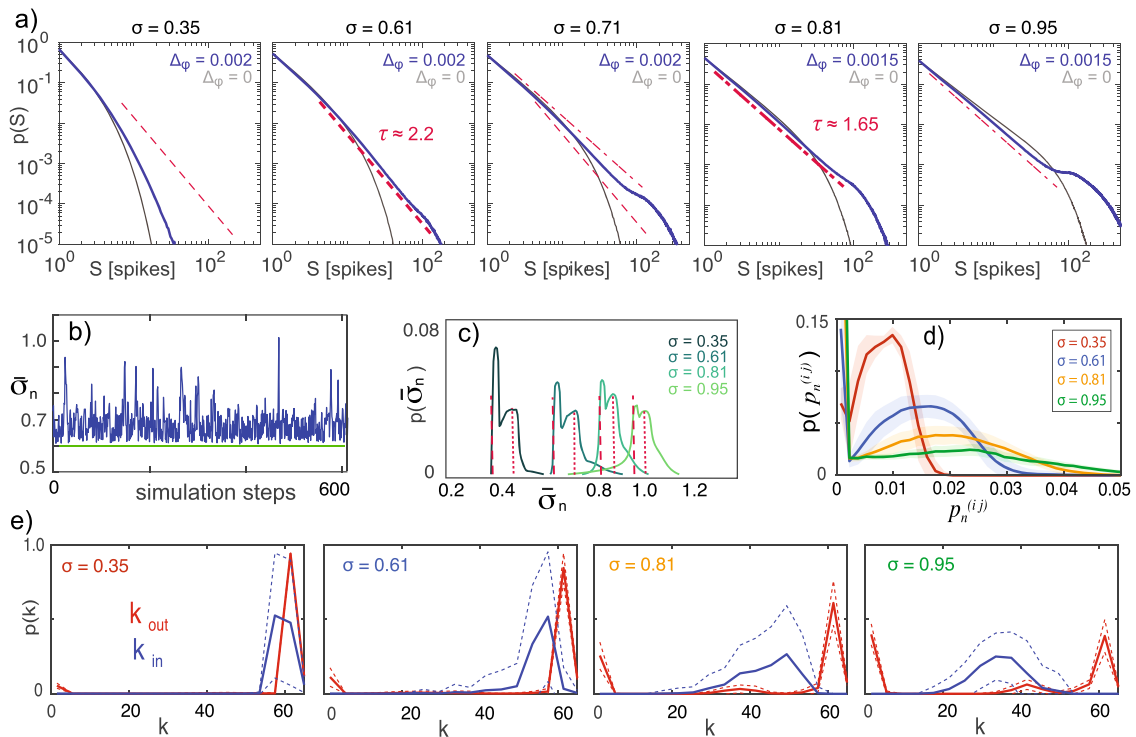


Fig. 1 Branching network model with facilitation exhibits signatures of different phase transitions. **a** Avalanche size distributions for the standard branching network with static (gray) and dynamic branching network with facilitation (blue) distributions, for the same branching parameter $\sigma \in \{0.35, 0.61, 0.71, 0.81, 0.95\}$. Dashed, dashed-dotted red lines are guides marking power laws $p(S) \sim S^{-\tau}$ for $\tau = 2.2$ and $\tau = 1.65$, respectively, highlighted where they describe the distribution well. **b** Temporal evolution of the network's effective branching parameter $\bar{\sigma}_n$ (blue) for $\sigma = 0.61$ (green line). **c** Distributions of $\bar{\sigma}_n$, indicating large fluctuations of the effective branching parameter. The red lines indicate presence of a second network component. **d, e** Change of the network topology with σ expressed by distributions of $p_n^{(i,j)}$, and of the degree k (where the in and out-degrees k_{in}, k_{out} are plotted separately). Colored areas and dashed lines, respectively, mark the one standard deviation boundaries ($N = 64$).

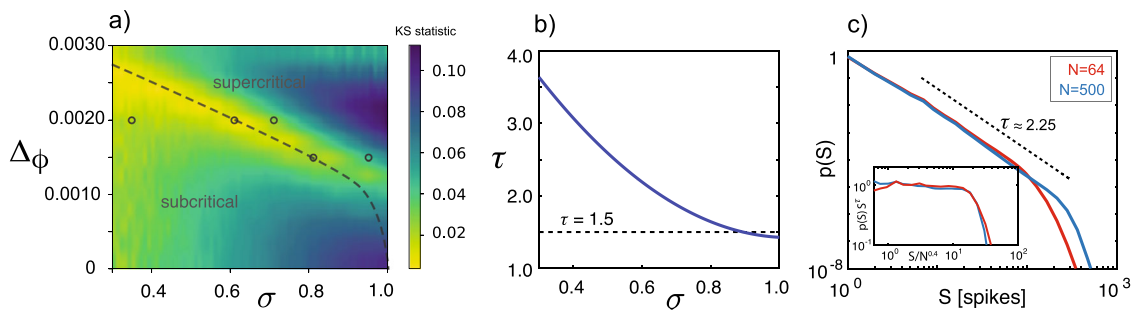


Fig. 2 Overview of the behavior of our dynamical facilitation branching network model (branching parameter σ and facilitation parameter Δ_ϕ). **a** Kolmogorov-Smirnov (KS) statistics of power law fits (smaller values: avalanche size distribution closer to a power law). Dashed line: Estimate of second order critical transition line. Circles: parameter values used to demonstrate early and late critical transitions. **b** Power-law exponent τ on the estimated second order critical line. Blue line: exponential curve through simulation data (perfect overlap, although obtained exponents slightly depend on the chosen fitting range). τ decreases for increasing σ and approaches for $\sigma \rightarrow 1$ and $\Delta_\phi \rightarrow 0$ the value 1.5. Dashed line: value of the classical critical branching model. **c** Scaling of our model with neural network size. Inset: finite size scaling collapse.

The general situation of how the two main modeling parameters (σ, Δ_ϕ) shape the generated avalanche size distributions support the hypothesis of a critical line (Fig. 2a), on which simulations exhibit optimal (in terms of the Kolmogorov-Smirnov (KS) statistics) power-law avalanche distributions with critical exponents τ decreasing upon the increase of σ (Fig. 2b). Notably, the figure embraces all the locations of the data presented in Fig. 1, including the suggested two critical regimes that fall on the line. Along the dashed line of suggested second order phase transitions, a good scaling behavior in system size is

observed, see, e.g., Fig. 2c). The two-humped distributions of the effective local branching parameters $\bar{\sigma}$ that within one network emerge across the parameter field displayed in Fig. 2a, may be at the origin of the line of critical regimes (see our Discussion section). The obtained results imply that in a network with facilitation, a transition from an early criticality regime (with a steep power law exponent) to a later late critical regime (with a milder power law exponent) can easily occur. For $\sigma \rightarrow 1$, the two humps merge into one hump (cf. Fig. 1c) and the difference to the original branching vanishes.

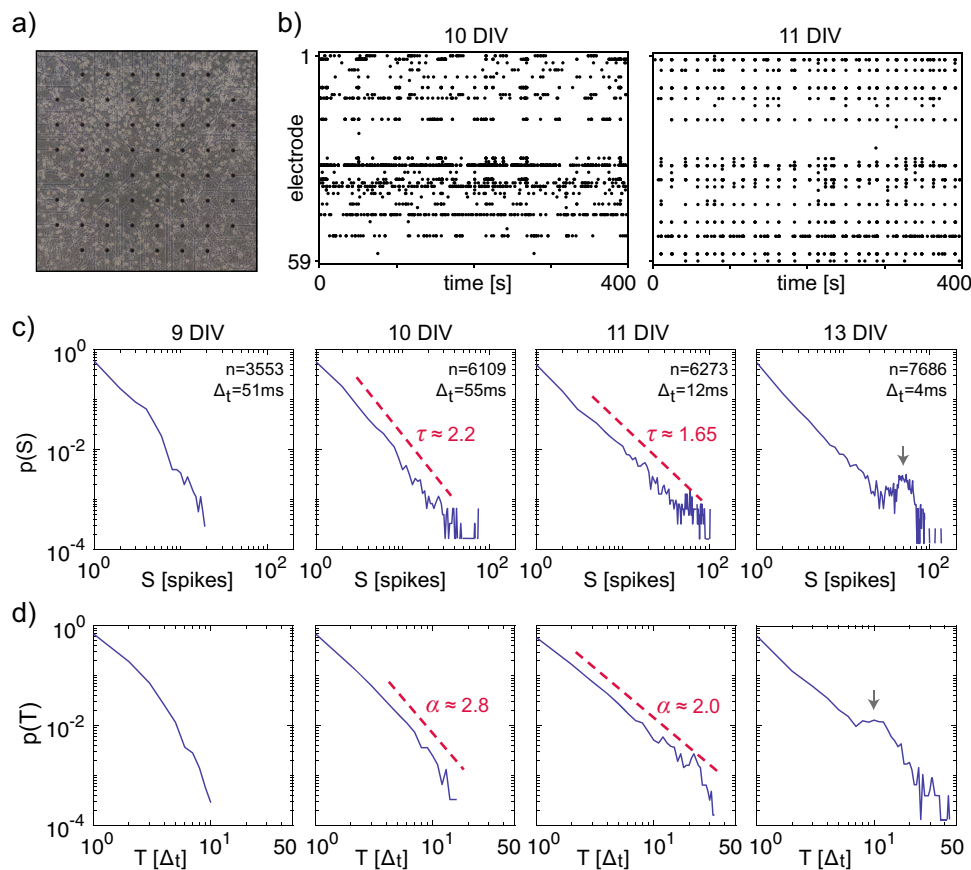


Fig. 3 Culture 1, overview: Transition between two avalanche criticality regimes during 10 and 11 days in vitro (DIV). **a** Fifty-nine-channel micro-electrode array (MEA) with plated hippocampal neurons. **b** Spike raster plots at 10 and 11 DIV (400 s out of 1 h recording). **c** Avalanche size S probability distributions $p(S)$. **d** Avalanche lifetime T probability distributions $p(T)$. Red dashed lines: statistically significant power law fits $p(S) \sim S^{-\tau}$ and $p(T) \sim T^{-\alpha}$. Gray arrows: humps indicating supercriticality. n : number of avalanches during daily recording time; Δ_t : temporal bin size used.

Experiments with two major regimes of criticality. In our experiments, dissociated hippocampal rat neurons were plated on a 59-channel micro-electrode array chip³³ (MEA) and cultured for several weeks (Fig. 3a, b). The density of plated cells influences how quickly the culture develops large activity bursts³⁴. We chose an intermediate level of $800 \text{ cells mm}^{-2}$, to be able to capture subtle day-to-day changes in neural activity. Starting from 9 days in vitro (DIV), a recording of 1 h of a culture’s spontaneous electrical activity was made every day, because of the extreme sensitivity of the cultures towards perturbations at this stage of the development. Neural avalanches are extracted from the (unsorted) spike recordings following the standard approach of binning time in discrete bins of width Δ_t related to the average firing rate of the network^{4,35} (see Methods section). An avalanche is then defined as the maximal extension of nonempty adjacent bins (i.e., each bin has at least one spike). Avalanche size S is defined as the number of spikes within the avalanches, and avalanche lifetime T is defined as the number of time bins in avalanche spans.

In these recordings we gather evidence that in the second week of in vitro development, a developing culture generally exhibits two distinct avalanche criticality transitions, within the course of 1–4 days. The two regimes—we again refer to them as the “early” and “late” criticality, due to their temporal occurrence—bear strong similarities to the two critical regimes exhibited above in our theoretical model and to what ref. ²⁸ found experimentally. Below we present the analysis of data from primarily two selected

cultures with conditioned media (“Cultures 1–2”, see our Supplementary Notes for additional cultures). Among the preparations without conditioned media, 3 out of 6 cultures showed very similar characteristics, at a generally temporally delayed development, with lower and less stable neural activity level expressions, and at a low level of astrocytes. In the brain, astrocytes together with other supporting (i.e., glial) cells are, however, at least as numerous as neurons³⁶, and they play a key role in the formation and functioning of synapses³⁷, suggesting that conditioned media cultures are closer to healthy in vivo tissue than unconditioned preparations. In the conditioned preparation, earliest manifestations of neuronal avalanches take place mostly in the second week of the development when the neuronal connections are still growing and strengthening³⁸ (and therefore the cultures cannot yet have reached the self-organized criticality regime characterizing “mature” ensembles²⁵). We observed that in this time window, the cultures pass through a variety of network phases, occasionally lingering in the close vicinity but not always evidencing a clear critical behavior. These changes happen very quickly, so that the capture of the characteristic illustrative time point was more difficult than in ref. ²⁵. Characteristic for the early period is also a bending of the histogram part for short events, which vanishes towards late criticality.

For the overall presentation of the observed general pathway we focus on Culture 1, before other data sets will exemplify important aspects of the prediction made by our model.

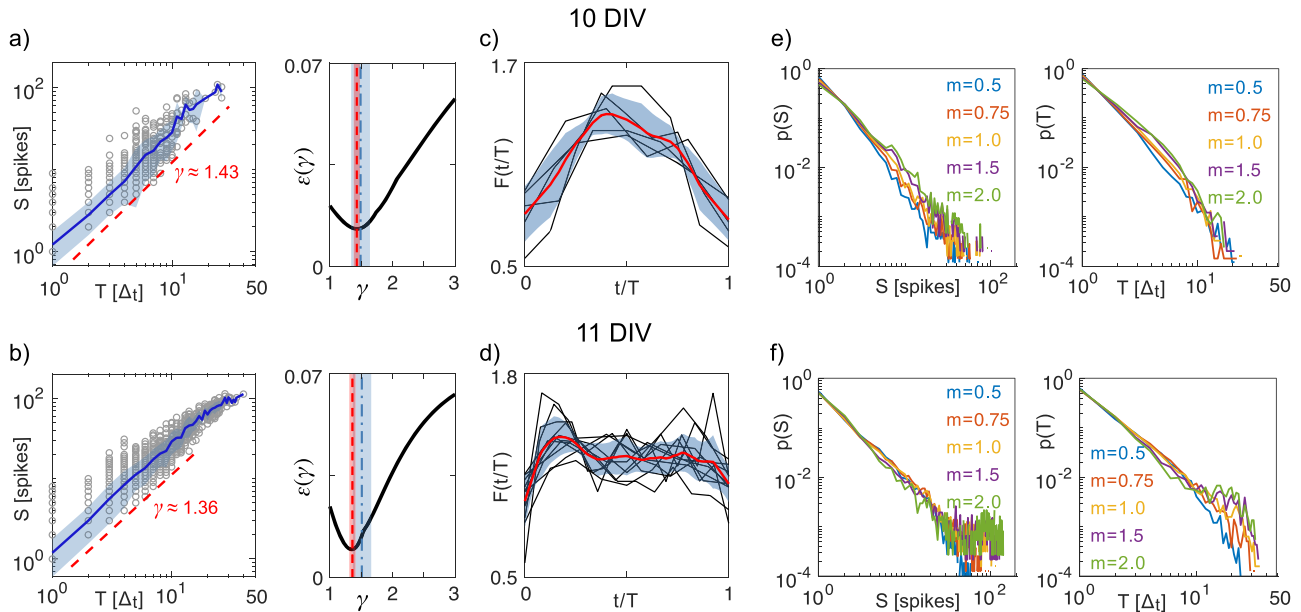


Fig. 4 Culture 1: avalanche criticality signatures, 10 days in vitro DIV (first row) and 11 DIV (second row). **a, b** Avalanche size S as a function of lifetime T (gray circles) and average size $\langle S \rangle(T)$ (blue lines). Red dashed lines: behavior predicted from the critical exponent γ (light-blue shaded area indicates one standard deviation from average), evaluated via γ_{\min} of the shape collapse error function ϵ in black, using $\langle S \rangle(T) \sim T^\gamma$ (blue dash-dotted line; γ_c estimated from crackling noise relationship). Light-blue shaded areas indicate one standard deviation from average. **c, d** Collapse of lifetime-rescaled average avalanche shapes, using γ_{\min} (black), and the average $F(t/T)$ (red); light-blue shaded areas indicate one standard deviation around average. **e, f** Avalanche size and lifetime distributions based on time bins of width $\Delta t_c = m\Delta t_c$.

This culture exhibits an exemplary case of two distinct avalanche criticality regimes occurring on two consecutive days of the development, see Fig. 3a, b. The interesting development starts at 9 DIV, where a subcritical phase with avalanches of smaller than system-size is exhibited (i.e., not all of the electrodes are activated), see Fig. 3c. At 10 DIV “system-size” avalanches appear (i.e., all electrodes are activated). Their appearance coincides with the emergence of scale-free avalanche activity characterized by a power-law exponent of the avalanche size distribution of $\tau = 2.18 \pm 0.05$ (obtained by the maximum likelihood approach^{39,40} fitted over the range $S \in [2, S_{\max}]$, with a p -value of $p = 0.13$; \pm refers to the 95% confidence interval; goodness-of-fit evaluation is based on the KS-statistic and the p -value has been calculated using 10,000 surrogate data sets; see Methods section). This critical regime is supposedly followed by a supercritical phase. For Culture 1, we missed the direct experimental data (but may infer this from the other two cultures analyzed), until on 11 DIV, the avalanche size distribution exhibits scale-freeness again, this time however, with a shallower exponent of $\tau = 1.65 \pm 0.05$ ($S \in [4, S_{\max}]$, $p = 0.30$). The new network critical regime is characterized by a more synchronous network firing (Fig. 3b). From 12 DIV onwards, the network enters a supercritical phase, highlighted by a “hump” in the size distribution near the value of the system-size (i.e., a bimodal distribution emerges), characteristic for a saturation process⁴¹. The avalanche lifetime distributions mirror this transition with power law exponents $\alpha = 2.76 \pm 0.16$ ($T \in [3, 15]$; $p = 0.12$ at 10 DIV), and $\alpha = 1.98 \pm 0.06$ ($T \in [2, 25]$; $p = 0.37$ at 11 DIV), see Fig. 3d. As can be expected from a distribution exhibiting a steeper decay, the scale-free behavior for $p(T)$ at 10 DIV is somewhat less striking than at 11 DIV. Generally, experimental lifetime power laws are often less satisfactory compared to their avalanche size counterparts^{4,6,7}. Seen in comparison to the cited results, the quality of the $p(T)$ power law at 11 DIV is, however, quite exceptional. The observed pathway suggests that we deal

with two separate phase transitions¹⁸. Such an interpretation will be corroborated by our analysis of Cultures 2 and 3 and our overview given at the end of the section.

To pin down avalanche criticality as the origin of the observed power law distributions (power law can be caused by several mechanisms⁴²), we use previously established tests^{13,43,44}, cf. Fig. 4. Critical phenomena suggest the presence of yet another power law describing the dependence of the average avalanche size on the lifetime^{43,45} with a third critical exponent γ as

$$\langle S \rangle(T) \sim T^\gamma. \quad (1)$$

Critical phenomena further imply that the mean temporal profiles of avalanches of different lifetimes should be similar across all scales, so that, when appropriately rescaled^{43,45}, collapse onto a universal scaling function

$$F(t/T) = T^{1-\gamma} s(t/T, T), \quad (2)$$

where $s(t/T, T)$ is the mean number of spikes at rescaled time t in an avalanche with the total lifetime T ($t = 1, 2, \dots, T$). In practice, the assessment of whether the (usually noisy) data exhibits a good enough shape collapse after rescaling may, however, be largely subjective, so that sometimes only a small subset of the data achieves a satisfying collapse⁴⁴. Therefore, we follow ref. ⁴⁴ that proposes to find a value of γ that minimizes the spread of the rescaled avalanche shapes, which can be objectively quantified by the shape collapse error $\epsilon(\gamma)$, i.e., $\gamma_{\min} = \operatorname{argmin}_\gamma \epsilon(\gamma)$ (see Methods section). This obtained value of γ is then tested as to whether Eq. (1) is satisfied. From theory, by the obtained exponents the crackling noise relationship⁴³

$$\gamma_c = \frac{\alpha - 1}{\tau - 1} \quad (3)$$

should be satisfied. If all three estimates of the critical exponent match, a suggested avalanche criticality is strongly

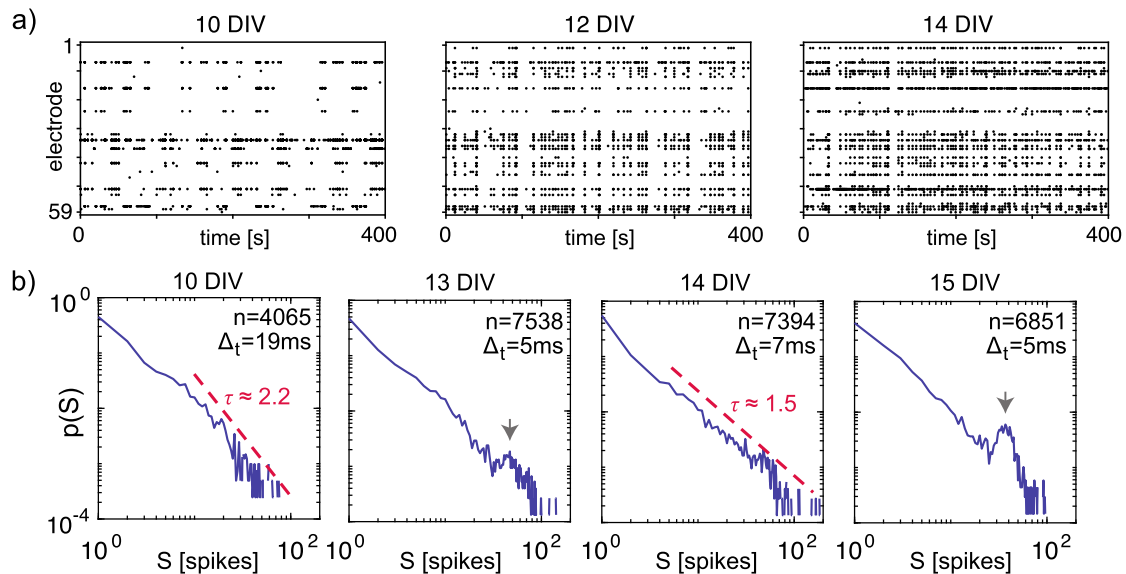


Fig. 5 Culture 2: a supercritical phase (11–13 days in vitro) separates an incompletely expressed early (10) from a late (14) days in vitro (DIV) avalanche criticality regime. a Spike raster plots demonstrating changed neural activity patterns. **b** Avalanche size distributions: Dashed lines are guides to the eye for statistically significant power law fits $p(S) \sim S^{-\tau}$. Gray arrows point to the so-called humps that indicate a supercritical phase. n : total number of avalanches on each day; Δ_t : temporal bin size.

corroborated. Our data from Culture 1 shows this to be indeed the case, with $\gamma = 1.43 \pm 0.05$ and $\gamma = 1.36 \pm 0.05$ at 10 and 11 DIV (Fig. 4a, b, respectively). At 10 DIV, we find that the shape collapse error is minimal for $\gamma_{\min} = 1.43$. The theoretical estimate from the crackling noise relationship yields $\gamma_c = 1.49 \pm 0.15$, and the relationship between average avalanche size and lifetime yields an estimate of $\gamma \approx 1.43$. The universal scaling function is parabolic and symmetric (Fig. 4c), congruent with observations in other experimental studies^{43,44,46}. At 11 DIV, the agreement between $\gamma \approx 1.36$ and $\gamma_{\min} = 1.35$ is also excellent, whereas the estimate $\gamma_c = 1.51 \pm 0.15$ slightly deviates from these values. Using 95% confidence intervals, the three estimates can, however, still be reconciled to be consistent. The rescaled avalanche is more noise-prone in this case, and $F(t/T)$ is flatter and more asymmetric (Fig. 4d). We suspect that temporal profiles of avalanches associated with shallower critical exponents will naturally be more variable (see our cluster analysis below for support), and thus we expect them to converge more slowly to the average shape, with respect to the number of samples. While to pin down the degree of asymmetry of the shape a larger number of samples would be required, previous experimental studies in dissociated neural cultures and other physical systems showing avalanches have also evidenced strongly asymmetric shapes^{43,45}. Lifetimes used for the collapses in Fig. 4c, d) span a range of 275 ms and 120 ms, respectively, covering a remarkable range (compared to 30–40 ms of earlier investigations^{43,46}).

Genuine scale-free behavior would imply that the choice of the temporal bin size Δ_t does not impair the power law property of avalanche size distributions¹³. The avalanche size and lifetime distributions at 10 DIV are only mildly affected by choosing different bin sizes $\tilde{\Delta}_t = m\Delta_t$ ($m \in \{0.5, 0.75, 1.0, 1.5, 2.0\}$), cf. Fig. 4e. At 11 DIV, the effect of choosing different bin-sizes is slightly stronger; a small hump emerges for significantly enlarged bin sizes for both size and lifetime distributions (Fig. 4f). This indicates that the recordings at 11 DIV are somewhat closer to the supercritical phase. Similar supercritical responses towards temporal binning have been found for avalanches in organotypic

cultures⁴³ exhibiting critical exponents very close to those at 11 DIV. The paradigm of transitions between two or more regimes of criticality seems to generally hold for our type of developing cultures; Supplementary Figs. 1–3 of the Supplementary Notes 1 and 2 contain corresponding verifications for two more cultures analyzed in detail. Culture 1 is our key example for the occurrence of two critical regimes; for the occurrence of a supercritical phase between the two we refer to Culture 2 (and to Culture 3 of the Supplementary Note 2). Admittedly, the captured evidence for early criticality for Culture 2 is weaker than for Culture 1; the collection of evidence seems, however, appropriate in the light of the cultures' quick development and our experimental constraints.

Figure 5a, b shows that at 10 DIV, Culture 2 can be interpreted to be close to the early critical regime, with $\tau \approx 2.23 \pm 0.13$ ($S \in [8, S_{\max}]$; $p = 0.12$) and $\alpha = 2.64 \pm 0.28$ ($T \in [5, 15]$, $p = 0.81$). Compared to Culture 1, exponent $\gamma = 1.28 \pm 0.03$ has a somewhat lower value (Fig. 4), but is in excellent agreement with the other two estimates $\gamma_c = 1.33 \pm 0.27$ and $\gamma_{\min} = 1.31$ (Supplementary Note 1, Supplementary Fig. 1). Over the next 3 days of development, the size distribution develops a prominent hump at large values of S , while the overall trend of the distribution remains consistent with the steep power law (avalanches at very short length scales carry often the effect of local behavior that is stronger than the collective effects at this scale⁴⁵). Compared to textbook examples, the humps look less impressive, but their size is in agreement with what we expect from related studies^{41,47} (cf. our Supplementary Note 2 for further discussion on the robustness of the supercritical humps). At 14 DIV, $p(S)$ transforms to a scale-free distribution consistent with late criticality, with exponent $\tau = 1.53 \pm 0.06$ ($S \in [6, 78]$; $p = 0.12$). The lifetime distribution is somewhat truncated, exhibiting reduced exponents $\alpha = 1.60 \pm 0.07$ ($T \in [2, 10]$; $p = 0.14$) and $\gamma = 1.19 \pm 0.05$, the latter value, however, still agreeing with $\gamma_{\min} = 1.13$ and $\gamma_c = 1.13 \pm 0.18$. At 15 DIV, the network is again in a supercritical phase. The spiking activity at 10 and 14 DIV (Fig. 5a) bears strong similarities to the patterns observed at early and late criticality, respectively (cf. Fig. 3b), whereas the supercritical phase between

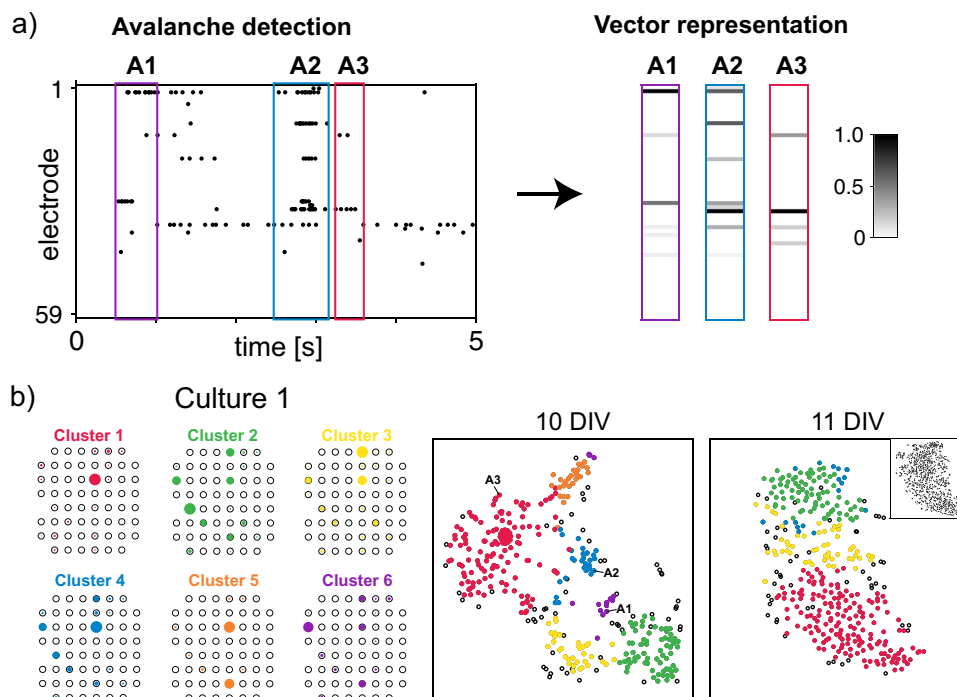


Fig. 6 Extraction of high-dimensional avalanche patterns. **a** Examples of detected spike avalanches with their vector representations, where the gray-scale indicates intensity. A1, A2, and A3 the three example avalanches extracted from Culture 1. **b** Each cluster's average avalanche vector is visualized by the sizes of the circles of the respective characteristic vector elements. For Culture 1, six clusters of avalanche patterns are found at early criticality. The clusters found at early and late avalanche criticality (10 resp. 11 days in vitro (DIV)) are displayed using a 2D-visualization tool (t-SNE embedding). Empty circles correspond to avalanches that were not assigned to a cluster. The inset at 11 DIV shows the embedding of all avalanches.

the two criticalities exhibits an elevated degree of synchrony and more prolonged periods of silence between activity bursts. Culture 3 corroborates our findings by exhibiting an example of two avalanche critical states being separated by a supercritical phase (Supplementary Fig. 2).

Reference¹⁸ highlighted that an avalanche critical point should be surrounded by a subcritical and supercritical phase. The recordings of Culture 2 (and of Culture 3 of the Supplementary Note 2) show that after the early critical point, a supercritical phase is entered (cf. Supplementary Note 2, Supplementary Fig. 2), to finally arrive at late criticality, followed by a supercritical phase (in Culture 1 we failed to catch the quickly occurring intermediate supercritical phase due to our experimental constraints). In this way, binning-independent power law distributions, universal scaling functions, and the verified fundamental relation between critical exponents, strongly suggest that roughly across two consecutive days, the neural cultures are in close vicinity of two distinct critical network regimes. Owing to their temporal succession, we will refer to the critical regimes the “early” and “late” criticality regime, respectively. While experimental evidence of avalanche critical regimes bracketed by subcritical and supercritical regimes has been provided earlier^{18,22}, the observation of such transitions linking distinct avalanche criticalities in a single preparation is, to the best of our knowledge, new.

Discussion

From the neuroscience viewpoint, our findings raise the question of whether the different regimes could be associated with distinct functional properties. While networks at criticality have been associated with enhanced computational power in the sense of a richer activity pattern repertoire and a greater flexibility of

responses to stimuli^{21,22}, it is not clear how such optimality arguments could be generalized to a line or series of critical regimes (see, however, refs. ^{17,20} for how criticality relates to computation). To distinguish between potential functional differences of early vs. late criticality, we performed a high-dimensional cluster analysis of the structure of neural avalanches in these regimes, using a vector representation that encodes the relative involvement of each electrode in an avalanche (Fig. 6a). Denoting by s_i the number of spikes registered during an avalanche at electrode i , we define an “avalanche vector” as $\mathbf{A} = [a_1, a_2, \dots, a_N]$, where $N = 59$ is the number of electrodes, and $a_i = s_i / \max\{s_1, s_2, \dots, s_N\}$. In this way, \mathbf{A} can be seen as a spatial encoding of the avalanche on the MEA (Fig. 6b), where the set of all avalanche vectors spans a 59-dimensional “feature space”. To see how the avalanches are organized in this space, we used an unsupervised Hebbian learning clustering (HLC) developed by us^{48–51}. This approach finds clusters of arbitrary shapes, without prior knowledge of the number of clusters or requiring a data dimensionality reduction step that generally distorts and biases the distances between data points⁵². Moreover, HLC has the capacity to filter out noise, by leaving such data without assignment to a cluster. Similar to other approaches, HLC uses an iterative optimization procedure that may depend on the initial condition, unless the clusters are clearly separated. To compensate for this, we ran the algorithm several (here, 15) times and accepted the highest number of clusters produced as the final result, which provides an upper bound for avalanche diversity estimates. To further quantify this estimate, we calculated a “cluster entropy” $H_c = -\sum_i c_i \log(c_i)$ based on the fraction c_i of all avalanches assigned to the i -th cluster. In this way, a small cluster entropy indicates that large clusters dominate the data, larger values indicate a more fair distribution. To ensure an objective comparison, across the DIV, the number of avalanche

samples used was kept the same for the estimate (except for 9 DIV that generally produced a much smaller number of avalanches; see Methods section). Fig. 6b) provides 2d-visualizations of the clustering results at early and late criticality. HLC finds consistently the highest number of distinct clusters at 10 DIV, when all cultures are at (or in a close vicinity to) early criticality (cf. Supplementary Note 3, “Cluster Analysis”, Supplementary Fig. 8). The subsequent supercritical phase exhibited a decrease in the number of clusters. Late avalanche criticality finds again a mild—but statistically significant—increased cluster number (without, however, reaching the level of early criticality), see Supplementary Note 3, Supplementary Fig. 9a). Our cluster entropy values (cf. Supplementary Note 3, Supplementary Fig. 9b) and low-dimensional visualizations corroborate the observation that early criticality offers the highest diversity of distinct avalanche patterns. To probe whether after early criticality, the avalanche patterns converge towards a smaller subset of patterns potentially due to a “learning” preference of a particular pattern, or whether clusters simply expand and merge (towards larger clusters of increased intra-cluster variability), we checked the median Euclidean distance d_{med} between the avalanche vectors. At late criticality, this value proved to be comparable or even larger than that at early criticality, cf. Supplementary Note 3, Supplementary Fig. 9c). At early criticality, the avalanche patterns are organized according to more separated clusters, in contrast to late criticality, where they appear to explore more variable patterns, having some boundaries between previously distinct clusters removed. Larger critical exponents seem, therefore, to indicate a preference of more local network interactions (as avalanches are less likely to involve the whole of the networks); more strongly connected cultures at a later development stage seem to join previously anatomically more separated subnetworks, enabling in this way more diverse spatiotemporal patterns. However, such interpretations need to be taken with care, as the relationship between cluster structures and network development is far from trivial. The results, however, point to different possible functional benefits offered by the two avalanche criticalities. The clearer clusters of early criticality appear to be more suited for a simple “state-coding”, where an avalanche cluster could be seen as a read-out codeword by an observer process. A similar coding strategy has been identified in weakly coupled periodic systems in terms of Arnold tongues⁵³, supporting efficient, Huffman-like input coding^{49,54}. The larger space spanned by the avalanche vectors at the late criticality, in contrast, is more advantageous for functions requiring more variable network responses, similar to earlier formulated optimality arguments regarding dynamic range, stimulus representation and information capacity at the avalanche criticality regime with $\tau \approx 1.5$ ²¹. This perspective suggests that the critical regime with steeper exponents may be associated with a stronger computational performance, in the sense of computation as a reduction of complexity of prediction⁵⁵. A similar interpretation was recently provided for the behavior of the peripheral auditory system where focusing on a particular sound was shown to tune the hearing sensor away from the initially unbiased critical regime with $\tau \approx 1.5$ ¹⁷. Comparative studies of the functional benefits of the different critical points promise to be an interesting topic of future research, not only for the goal of basic understanding of nervous systems, but also for augmented machine learning efficiency based on critical dynamics^{56,57}.

Early and late criticality bear an intriguing similarity to the two different universality classes for cultured neuronal networks found by Yaghoubi et al.²⁸. In their study, the two universality classes were found in two different types of neuronal cultures that either had or didn't have an acid metabolite 5M4Hfolate in the culture medium. The presence of folate was conjectured to affect

the structural network connectivity and lead to the observed change in critical exponents from $\tau = 1.65 \pm 0.1$, $\alpha = 2.15 \pm 0.2$ and $\gamma = 2 \pm 0.4$ in their “control” cultures, to $\tau = 2.2 \pm 0.2$, $\alpha = 3.3 \pm 0.4$ and $\gamma = 2 \pm 0.4$ in the cultures with folic acid. Both sets of exponents correspond rather well to the exponents we find at the early and late criticality: The agreement is perfect for τ , and α is also fairly consistent within the 95% confidence intervals. The discrepancy in the values obtained for γ can be explained as follows. In corresponding simulations, γ has been reported to be particularly sensitive to the network structure⁴³, and to approach $\gamma = 2.0$ for all-to-all connectivity. Since the recordings by Yaghoubi et al. were made in the third and fourth weeks of development at denser connectivity³⁸ (compared to the 10–11 DIV period of our measurements), differences in the estimates of γ may be expected. Finally, the large difference in Δ_t between the early and late criticality is also consistent with the experiments of ref. ²⁸. For an extended discussion, including other related works, we refer the reader to Supplementary Note 4 (“Discussion of the critical line paradigm”).

In our modeling and experimental work we revealed that more than one single critical regime can occur, and that thus neural network criticality may be richer and more complex than is generally anticipated. A cluster analysis of *in vivo* avalanche patterns revealed that the different criticality states likely support different roles of information processing and computation. While regimes with smaller power-law exponents suggested in ref. ²¹ seem to more support the information representation aspect, regimes with the additionally detected critical regime with a stronger exponent has a stronger connection to computation, interpreted as the simplification, i.e., destruction of information⁵⁵. It is this link that associates different functional roles to the two observed criticalities. In the context of what we call here “late” criticality, the importance of dynamically changing network interactions for avalanche criticality has been pointed out⁵⁸. Here, we have extended this insight by demonstrating that facilitation may be at the basis of other experimentally observed avalanche-critical regimes as well. Often, measurements that deviate from the expectations of an established theory can be explained away by not ideal experimental or computational conditions. However, in some situations exactly the deviations might provide the key for understanding the underlying system⁴¹. Rather than as a correction of the concept, we see the presented results in the perspective of offering a constructive generalization of the fruitful application of the physics concept of criticality, towards new observations in biological neural networks and, probably to biology beyond. As neuroscience experiences a technological revolution and better tools for recording and manipulating neural activity in the brain are becoming available, further experimental investigations will be able to probe the existence, origin and functional properties of distinct avalanche criticality states *in vivo*.

Methods

A dynamic branching model with facilitation and depression. In many works of the past, the existence of a single critical point seems to have been favored, implicitly or explicitly. However, fundamental examples of statistical physics models demonstrate that critical regimes depend on the dimension of the underlying topology. Upon neuronal development, this topology—while in most cases nontrivial to grasp—must be expected to change. If we can trust the analogy to statistical physics, it is thus not impossible to see the emergence of various critical regimes during the development of a neural culture. In our contribution we scrutinize the question whether such an occurrence is likely to emerge, jointly from the theoretical, modeling and experimental sides.

In particular during the early neural culture maturation process, the wiring structure among the neurons undergoes big changes. Developing neuronal cultures provide an opportunity to observe how synaptic connection density and strength gradually increase over the first couple of weeks³⁸. The interaction strength between the neurons in the culture is naturally “tuned” from weaker to stronger,

across consecutive days in vitro (DIV) early development (before synaptic pruning and effects of inhibition become prominent²⁵). DIV can be seen as a proxy for a “control” or “tuning” parameter of the culture, which enables the observation and analysis of different, naturally occurring network activity phases. In our experimental data, we find two critical regimes.

While the evidence we provide will suffer from all the typical shortcomings of this concept in biological applications, theoretical arguments and modeling results strongly support our experimental observations. From a theoretical model of networks based on a “rich get richer with a local constraint” principle⁴¹, we already know that abstract biological systems can, in principle, converge to criticality anywhere on a line of critical states. The simplest manifestation of this model are compounds that tend to aggregate according to the rich-get-richer paradigm (e.g., due to a global attraction scheme such as gravitational, or electrical attraction), but where this growth is limited by local constraints (such as a finite number of docking sites to the compound). The paradigm generates, apart of the critical line, regimes characterized by humped distributions where the local constraint is incompatible with the global principle (for details see ref. 41). In loose physics terms, the humps can be interpreted as fingerprints of a “friction” that becomes apparent at a particular scale. In these systems—that are by construction non-critical—the humps in the distributions may be apparent or barely visible, but where they are apparent, their form strongly agrees with what we observe in our experimental data. Below, we will use this model to interpret the structural changes seen in the development of our experimental cultures during maturation. A recent model of elementary adaptive network automata supports our view⁴⁷. Our method to detect criticality does not rely on the goodness of fit for power laws (that we of course consider as important first indicators). To identify criticality (in model simulations as well as in experiments) one needs to demonstrate that the system is spatially scale-free, i.e., avalanches are required to follow a power law (1), that it is temporally scale-free (2) (for this, different binnings should be used that also should exhibit a power law relation) and that it is space-time scale-free (3) (for this, the existence of a scaling function should be evidenced, using extracted spatial and temporal information). Finally (4), the consistence of all these tests should be evidenced by showing that the crackling noise relationship among the estimated power-law exponents is satisfied with reasonable accuracy.

That the emergence of one single critical regime may not be a necessity is, moreover, emphasized by the statistical thermodynamics of alloys: If contributing compounds have sufficiently different characteristics, lines of critical regimes emerge⁵⁹. In biological neural networks and typical modeling approaches, different components are present, too, at different levels. Normally, excitatory and inhibitory components are too strongly interwoven to become relevant in this context. However, if the well-known neuronal properties of neuronal facilitation and depression are included into the classical avalanche branching model, subpopulations (consisting each of excitatory and inhibitory neurons) emerge that express very distinct dynamical and topological properties. We blame it on this that in our simulations a line of critical regimes emerges that seems to be approached by our experimental cultures in variety of locations during their development (generating distinct critical exponents).

Our modeling approach is a refinement of an accepted theoretical model for explaining criticality in neural networks (producing a single regime of criticality). To arrive at an analog of the structural and behavioral changes that our experimental neural network cultures undergo during maturation, we complete this model with the inclusion of facilitation and depression, basic mechanisms at work in neural networks. From this model we obtain evidence for the existence of a line of critical regimes, which corroborates our observation of at least two main avalanche criticality regimes in our experiments with maturing dissociated neural cultures.

Computational models exhibiting avalanche criticality regimes in neural networks have mostly focused on explaining self-organization aspects of the “late” criticality regime characterized by $\tau = 1.5^{22,58,60,61}$. Only in few, mostly numerical, studies an “early” criticality regime characterized by $\tau \approx 2.0^{58}$ and $\tau \approx 2.5^{20,62,63}$ was found, essentially without explaining the origins of the difference of these regimes (we use the terms “early” and “late” relative to their occurrence in the experimental setting on the time axis of development). By demonstrating a simple toy model that only uses basic biological activity propagation assumptions that gives rise to avalanche size distributions consistent with both, the early and the late, regimes of criticality, we fill this gap. We start from the network model of ref. 22, describing a branching process²³ that we map on a directed network at a mesoscopic scale. In this setting, each experimental electrode can be viewed as being represented by a network node. Such a node is either active (i.e., spiking) or silent. Active nodes i excite stochastically other network nodes j , where the propagation of activity is according to probabilities $p^{(ij)}$. The sum of all outgoing excitation probabilities of each node is constrained to a value σ , the so-called “branching parameter” that determines the expected number of new active nodes (also known as “descendants”) produced by a single active node (more details see below). For $\sigma = 1$, this model yields power-law distributed avalanche sizes with the scaling exponent $\tau \approx 1.5^{22}$. For $\sigma > 1$, the network is in a supercritical phase, whereas for $\sigma < 1$ the size distribution becomes subcritical with the probability of larger avalanches decaying rapidly (Fig. 1a), in gray). Critical states with power laws of exponent $\tau \approx 2.2$ are not generated by this model.

The branching network is a minimal model for some neural avalanche dynamics measured on micro-electrode arrays (MEA). Following ref. 22, each electrode i , $i = 1, 2, \dots, N$, is represented by a “binary processing unit” $b_n^{(i)}$ that either can be active ($b_n^{(i)} = 1$) or inactive ($b_n^{(i)} = 0$), where n denotes the simulation time step. Each unit can be randomly connected to N_C other units; we consider $N_C = N - 1$, i.e., all-to-all connectivity²², which permits functional connections between all MEA electrodes.

Connection from unit i to unit j is defined by an activation (excitation) probability $p^{(ij)}$. This probability describes the likelihood of unit j becoming active if unit i was active in the previous time step. The activation probabilities for each connection are chosen randomly, but the sum of each units’ outgoing activation probabilities is constrained to $\sigma^{(i)} \equiv \sum_{j=1}^{N_C} p^{(ij)} = \sigma$. σ is network’s branching parameter, i.e., the expected number of active descendants that an active unit produces. More formally, the activation of unit j by unit i is determined by a binary ‘transmission’ variable

$$t_n^{(ij)} = \begin{cases} 1, & b_n^{(i)} = 1 \wedge r < p^{(ij)}, \\ 0, & b_n^{(i)} = 0 \vee r \geq p^{(ij)}, \end{cases} \quad (4)$$

where r is a random number drawn uniformly from the interval $(0, 1)$. Units update their state each time step; they become activated if any of their incoming connections had a transmission of value 1 that did not fall into the refractory period of length $t_R = 2$ simulation steps:

$$b_{n+1}^{(i)} = \begin{cases} 1, & \sum_{k=1}^{N_C} t_n^{(ki)} > 0 \wedge b_m^{(i)} = 0, m \in (n - t_R, n], \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

For exciting the units, several approaches can be followed²². We chose the one that does not require any new free parameters, by exciting one randomly chosen unit whenever an avalanche has finished (i.e., when $b_n^{(i)} = 0$, for all i). Because of $\sigma_i = \sigma$, this setting is referred to it as the ‘static’ branching network model.

An addition of dynamic facilitation $\phi_n^{(i)}$ and depression $\delta_n^{(i)}$ brings the standard model closer to our experimental data. Earlier experiments have shown that during in vitro neuronal maturation, both of these opposing effects are observed: The growth in the network connectivity, measured in terms of, e.g., synaptic density and size, number of synapses per neuron and number of synaptic vesicles per synapse³⁸, as well as a concurrent decrease in neuronal excitability in terms of, e.g., a decreased resting membrane potential^{29,30}. These effects have been reported for different types of experimental neuronal culture preparations and they appear to be general features of neuronal maturation. The development of connectivity and excitability is brought about by both genetically predetermined and activity-dependent mechanisms^{64,65}, while the opposing developmental effects might support activity level homeostasis⁶⁶. We have attempted to compactly express the connectivity and excitability properties using mainly the branching parameter σ and facilitation Δ_ϕ . As outlined above, biological experiments suggest that when modeling neuronal development, σ should be gradually increased, whereas Δ_ϕ ought to be decreased.

Facilitation variable $\phi_n^{(i)}$ is incremented by Δ_ϕ if a connection from unit k to unit i fails to transmit, and otherwise suffers from an exponential decay; if unit i becomes activated, $\phi_n^{(i)}$ is reset to 0:

$$\phi_{n+1}^{(i)} = \begin{cases} 0, & b_m^{(i)} = 1, m \in (n - t_R, n], \\ \eta_\phi \phi_n^{(i)} + \Delta_\phi \sum_{k=1}^{N_C} (1 - t^{(ki)}) b_n^{(k)}, & \text{otherwise.} \end{cases} \quad (6)$$

For the whole duration of the refractory period, $\phi_n^{(i)} = 0$. A depression $\delta_n^{(i)}$ is added that increments after unit i has been active to counter-balance potentially emergent runaway activity:

$$\delta_{n+1}^{(i)} = \eta_\delta \delta_n^{(i)} + \begin{cases} \Delta_\delta, & b_n^{(i)} = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Facilitation and depression turn the activation probabilities into dynamic variables:

$$p_n^{(ij)} = p^{(ij)} + \phi_n^{(j)} - \delta_n^{(j)}, \quad (8)$$

where $p^{(ij)}$ is the baseline excitation probability, constrained by the initially set σ . The dynamic excitation probabilities replace the static probabilities of Eq. (4) as

$$t_n^{(ij)} = \begin{cases} 1, & b_n^{(i)} = 1 \wedge r < p_n^{(ij)}, \\ 0, & b_n^{(i)} = 0 \vee r \geq p_n^{(ij)}. \end{cases} \quad (9)$$

From a complex systems perspective, the branching parameter σ can be seen as a “global” network-level property that sets the limits of network connectivity, whereas facilitation is a “local” individual network node property that modifies on top of the connectivity “backbone” defined by the branching parameter the local activity, which generates the in vitro observed activity properties.

We refer to this modified model as the “dynamic” branching network, since the effective branching parameter of unit i , $\sigma_n^{(i)} = \sum_{j=1}^{N_C} p_n^{(ij)}$ is no longer constrained to a fixed σ , but we may determine the instantaneous branching parameter by taking the average across all units, $\bar{\sigma}_n = 1/N \sum_i \sigma_n^{(i)}$. The model parameters used to obtain

the results shown in Fig. 1a–e are for the static network, $\Delta_\phi = \eta_\phi = \Delta_\delta = \eta_\delta = 0$. For the dynamic model, we chose $\Delta_\phi = 0.002$ for $\sigma \in \{0.35, 0.61, 0.71\}$, $\Delta_\phi = 0.0015$ for $\sigma \in \{0.81, 0.95\}$, and $\eta_\phi = 0.35$, $\Delta_\delta = 0.15$, $\eta_\delta = 0.35$ for all σ . For the comparison we used a network size of $N = 64$ and a refractory period of $t_R = 2$ iteration steps in both cases; for each σ , averages from ten simulations of 10^7 time steps with randomly initiated $p^{(i)}$ are shown. Across the different simulations, we observed a negligible variation among the avalanche size distributions only. For $\sigma < 1$, avalanche size distributions with a steeper power-law can be expected, if activation failure is compensated by mild excitation facilitation: If an activated unit i does not succeed in exciting neighbor j the excitation probabilities $p^{(kj)}$ of unit j are increased by an amount Δ_ϕ . This facilitation is counter-balanced by an exponential decay in time, that in the absence of continued facilitation sends the excitation probabilities back to their baseline values σ (Fig. 1a, b). Such a behavior is suggested by biology: Subthreshold synaptic input to neurons is known to increase—transiently—neuronal excitability, either by reducing the distance between neuron's membrane potential and spiking threshold, or by short-term synaptic facilitation⁶⁷. When added to the original model, the effect is a more heavy-tail like decay of avalanche distributions. To prevent a dominance by very large avalanches, resource depletion³¹ that depresses the outgoing excitation probabilities of active units should naturally be added. In our study, depletion affected the statistics of very large avalanches only, so that facilitation emerged as the key mechanism for reproducing the experimental observations.

Neuronal culture preparations. Hippocampi were micro-surgically separated from E18 (embryonic day 18) Sprague-Dawley (SD) rat (Koatech, Republic of Korea). Dissected tissues were dissociated with pipetting in Hank's buffer salt solution (HBSS) and centrifuged at 1000 rpm for 2 min. Supernatant was then removed, and settled cell pellet was resuspended in plating medium (conditioned medium, explained below, supplemented with 12.5 μM l-glutamate). After being sieved through a cell strainer (BD Falcon, NJ, USA), cells were plated on a chip at the density of 800 cells mm^{-2} . For producing conditioned media, we first made astrocyte cultures by incubating dissociated culture of E18 SD rat cortex on poly-D-lysine coated glass in Neurobasal medium supplemented with 2 mM GlutaMAX, 10% horse serum, and 1% penicillin-streptomycin for 20 days. At 20 DIV, the whole media was exchanged into the maintenance media (Neurobasal medium supplemented with B27, 2 mM GlutaMAX, and penicillin-streptomycin) after washing the culture with maintenance media once. The conditioned media was collected after 3 days from the media change.

Cultured neurons were kept in a humidified incubator maintained at 37 °C and 5% CO_2 . Half of the medium was changed twice a week with conditioned medium. All experiments were performed in accordance with the guidance of the Institutional Animal Care and Use Committee (IACUC) of the Korea Advanced Institute of Science and Technology (KAIST), and all experimental protocols were approved by IACUC of KAIST.

Neural activity recordings. For the recordings, a micro-electrode array (MEA) from Multi Channel Systems (Reutlingen, Germany) with 59 microelectrodes (TiN, 30 μm in diameter, 200 μm spacing) and 1 reference electrode was used. To promote cell adhesion on chip surface, the MEA was coated with poly-D-lysine (100 $\mu\text{g}/\text{mL}$) and sterilized with 70% ethanol. To prevent the evaporation of cell culture medium, the MEA was capped with a FEP-membrane sealed Teflon ring. The MEA was placed on the heating plate which is connected to TC01 (Multi Channel Systems, Reutlingen, Germany) for keeping the temperature at 37 °C. A custom-built amplifier with the gain of 1000 and the bandwidth 150 Hz–4.5 kHz was connected to the MEA, and the amplified signal was digitized at 25 kHz with USB-ME64 system (Multi Channel Systems). For spike detection, the threshold level was set to six times the standard deviation of the background noise. Starting from 9 DIV, every day electrical activity of 1 h was recorded. To extract spike timestamp for the data analysis, NeuroExplorer (Nex Technologies, Madison, USA) was used. First 10 min of the recordings were discarded from further analysis, to avoid artifacts due to possible non-stationarities. The recording was performed after at least 12 h from the media exchange in order to avoid the perturbation effect.

Data analysis details. In the following subsection, symbol τ denotes the cross-correlation lag, in contrast to the rest of the paper, where it stands for the avalanche-size critical exponent. Neuronal avalanche extraction followed the conventional approach^{4,35} to bin time in discrete bins of width Δ_t and to define an avalanche as the maximal extension of nonempty (i.e., having at least one spike) adjacent bins. The value of Δ_t is commonly chosen equal to the average inter-event interval, (IEI), where IEI is the time difference between two consecutive events on the array. Events in our case were spikes registered at any of the recording array's electrodes (in other approaches, events can also be peaks in the local field potential (LFP) signal instead⁴; the use of LFP peaks for avalanche analysis, however, has recently been discouraged¹⁹). In our data, we often observed heavy-tailed IEI distributions leading to a divergence of $\langle \text{IEI} \rangle$ and thus also of Δ_t (cf. Fig. 7a–c). The IEI at the tail of the distribution relate to periods of prolonged silence, usually after bursts of elevated neural activity (Fig. 7a). In previous studies^{4,35}, this issue has

been dealt with by discarding the long inter-event intervals, i.e., Δ_t is set equal to an average $\langle \text{IEI} \rangle^*$ obtained by using only $\text{IEI} \in [0, \text{IEI}_{\text{max}})$. The cut-off IEI_{max} is defined as the value where the average pairwise cross-correlation between electrode LFP signals becomes negligible^{4,35}. We adapt this methodology for our case of spike signals, by calculating the average spike train cross-correlation⁶⁸ $C(\tau)$ across all pairs of electrodes (where only in this section, symbol τ codes for the cross-correlation lag), and setting IEI_{max} equal to τ at which $C(\tau)$ crosses the abscissa (Fig. 7d). We define a spike train at electrode i as $S_i = \{t_1^{(i)}, t_2^{(i)}, \dots, t_{N_i}^{(i)}\}$, where $t_n^{(i)}$ is the time when spike n occurred, $t_n^{(i)} < t_{n+1}^{(i)}$, and N_i is the total number of spikes registered by electrode i . The cross-correlation function between two spike trains S_i and S_j is approximated by a histogram $C_{ij}(\tau)$; τ takes discrete values $k\delta t$, where δt is the bin size of the histogram and $k \in \mathbb{Z}$. For computational efficiency, we consider only $\tau \in [-10^3, 10^3]$ ms. For binning we use $\delta t = 25$ ms, but the particular choice of δt does not influence the results (Fig. 7e). $C_{ij}(\tau)$ takes the value of the total number of times that spikes in S_i are separated in time from spikes in S_j by an interval $(t_x^{(i)} - t_y^{(j)}) \in [\tau - 0.5\delta t, \tau + 0.5\delta t)$, where $x \in \{1, 2, \dots, N_i\}$ and $y \in \{1, 2, \dots, N_j\}$. $C_{ij}(\tau)$ is further normalized by subtracting $N\delta t/(2T)$, where N is the total number of spike pairs considered. The normalization term expresses the expected value of $C_{ij}(\tau)$ if the intervals between the spike pairs were to be uniformly distributed. Finally, we calculate the average across all normalized pairwise cross-correlation functions, $C(\tau) = \langle C_{ij}(\tau) \rangle_{i,j}$, and find the zero-crossing, τ_0 , of $C(\tau)$ by locating the first bin where $C(\tau) < 0$ (Fig. 7d). Setting $\text{IEI}_{\text{max}} = \tau_0$ leads to exclusion of humps at the tail of IEI distributions (Fig. 7b), and prevents a divergence of $\langle \text{IEI} \rangle^*$ (Fig. 7c).

Avalanche size and lifetime distributions are estimated by maximum likelihood (ML) fits, following the guidelines in refs. ^{39,40}. The fits are discrete truncated power law distributions

$$p(x) = \frac{x^{-\theta}}{\sum_{y=a}^x y^{-\theta}}, \quad (10)$$

where $x = S$, $\theta = \tau$ is used for avalanche sizes and $x = T$, $\theta = \alpha$ for avalanche lifetimes. The range of a fit $[a, b]$ is determined as the longest interval for which the goodness-of-fit measure is still satisfactory (see below) and b is allowed to be no larger than the largest observed value of x . In cases when the empirical distribution exhibits a sudden and steep fall-off at the tail, b is limited by the location of the fall-off (e.g. $p(T)$ at 11 DIV in Fig. 3). For the evaluation of the goodness-of-fit, Monte Carlo simulations of the fitting process are performed⁴⁰. Ten thousand surrogate datasets using the ML estimate of θ are first generated, with the same number of samples as the original data used for fitting. The surrogates are then fit using ML, and the KS-statistic is calculated for each fit. The p -value p is calculated as the fraction of surrogate datasets with a higher KS statistic than the corresponding experimental data fit. A rather stringent criterion of $p > 0.10$ is set to consider a fit satisfactory.

To estimate the third critical exponent γ from the relation $\langle S \rangle(T) \sim T^\gamma$, we use the conventional least squares fitting over the maximal available range of T (unless a clear deviation from a straight line behavior is observed). Confidence intervals for the critical exponent estimates are found by using the “bootstrapping” method^{39,69}. Ten thousand surrogate datasets are generated by sampling uniformly at random with replacement from the original data, and the fitting procedure over the same range $[a, b]$ is repeated for each surrogate dataset. Two standard deviations of the resulting distribution of the parameter estimates yield the reported 95% confidence intervals. Alternatively, the confidence interval of the ML estimate is obtained via the Monte Carlo simulations described above, which we found to give fully consistent results if compared with the bootstrapping method.

Using the methods described above, we tested the hypothesis that there could be a critical transition at a particular day in development (i.e., at days between subcritical and supercritical phases, when the avalanche distributions exhibit a scale-free trend). Competing statistical distribution hypotheses to explain the distribution shape were not considered. We feel that such a comparison would reach beyond of what would be compatible with the sampling protocol that we had to use.

For testing avalanche shape collapse, we used the methodology introduced in ref. ⁴⁴. This improvement of an earlier approach⁴³ allowed us to use a much higher number of avalanches and to define an objective shape collapse criterion. To determine the quality of the avalanche shape collapse, the averaged and rescaled avalanche profiles of different lifetimes T , $F(t/T) = T^{1-\gamma} s(t/T, T)$ are first linearly interpolated at $n_t = 1000$ points along the scaled duration. The variance across the different $F(t/T)$ is calculated at each interpolated point, and the shape collapse error $\epsilon(t)$ is then defined as the mean variance divided by the squared span of the avalanche shapes, where the span equals the maximum minus the minimum value of all rescaled avalanche profiles⁴⁴. In the presented analysis, avalanche shapes of $T > 4\Delta_t$ with at least 20 samples were used, to avoid too noisy average temporal profiles.

Hebbian learning clustering^{48–51} (HLC) recasts the data (in our application: the avalanche vectors) as the nodes on a k -nearest neighbor graph (using Euclidean distance). The distances between data points furnish the initial edge weights of the graph. The nodes are represented by neurons that interact by spikes, mediated by the edge weight as the synaptic strength being dynamically updated following the

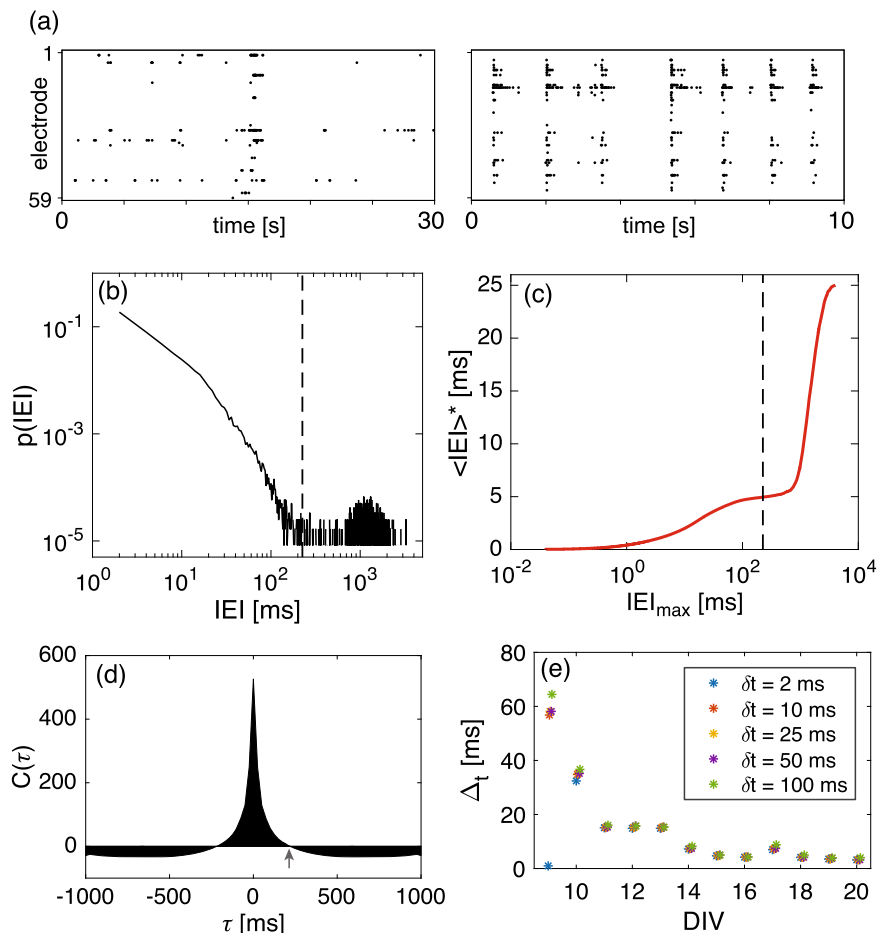


Fig. 7 Determination of the temporal bin size Δ_t . **a** Exemplary raster plots exhibiting typical network activity associated with heavy-tailed inter-event interval (IEI) distributions. **b** Example of a bimodal IEI distribution (solid curve) exhibiting a hump at the distribution's tail; dashed line marks the cut-off value $IEI_{\max} = \tau_0$. **c** A running average of IEI that is calculated by considering only $IEI \in [0, IEI_{\max})$ diverges as the cut-off IEI_{\max} increases (solid curve); dashed line marks $IEI_{\max} = \tau_0$. **d** Average, pairwise cross-correlation ($\delta t = 25$ ms); arrow marks zero crossing at $\tau_0 = 225$ ms. **e** Calculations of Δ_t for different days of development exhibit robustness with respect to the cross-correlation histogram bin size δt . Variation is largest for a small numbers of spikes (i.e., at 9 days in vitro).

Hebbian learning paradigm of neuroscience. Clusters emerge as synchronous groups of nodes with enhanced within-cluster edges and weakened—or altogether pruned—between-cluster edges. HLC is fully based on local k -nearest neighbor information and thus allows for an unbiased treatment of data (no prior information, e.g., number of clusters or assumptions on their shape, is needed). We use $k = 30$, which has been shown to perform well on comparable synthetic datasets⁵².

To ensure that the avalanche patterns represent the collective behavior of the underlying neural network, we restricted ourselves to avalanches with $S \geq 8$, for all cultures. At 9 DIV, for the cluster analysis, all avalanches were used (83, 189, and 70 avalanches for Cultures 1–3, respectively). Later, to ensure a fair comparison between different days in development, the number of data points used in the cluster analysis n_A was kept the number observed at 10 DIV ($n_A = 371$ for Culture 1, $n_A = 712$ for Culture 2 and $n_A = 439$ for Culture 3). In the following DIV that produced an increased number of events, the avalanches for the cluster analysis were subsampled (uniformly at random) from the pool of all available avalanches and the average of ten subsamplings were used for the final result. The subsampled datasets represent the full datasets sufficiently well. For the quantitative analysis, only clusters with ten or more avalanches were considered.

The organization of the avalanche vectors in the high-dimensional feature space was visualized by 2D-projections of the clustering results, using the t-distributed stochastic neighbor embedding⁷⁰ (t-SNE) algorithm. The perplexity parameter was set to 30; we show the best embedding (i.e., that with the smallest loss) out of ten simulation runs. While t-SNE can be quite powerful in revealing structures within the data, for complex cluster shapes it can also provide misleading results⁵². Hence the embedding must be interpreted with care: In Fig. 6b, a few points that are separated in the embedding by t-SNE, should be in the same group, according to HLC that is more trustful. Our reported observations were, however, tested to not be impaired by such effects.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets as well as the methods used in the present study are available from the corresponding authors on reasonable request.

Received: 4 December 2018; Accepted: 17 December 2019;

Published online: 17 January 2020

References

- Vogels, T. P., Rajan, K. & Abbott, L. F. Neural network dynamics. *Annu. Rev. Neurosci.* **28**, 357–376 (2005).
- Sussillo, D. Neural circuits as computational dynamical systems. *Curr. Opin. Neurobiol.* **25**, 156–163 (2014).
- Ringach, D. L. Spontaneous and driven cortical activity: implications for computation. *Curr. Opin. Neurobiol.* **19**, 439–444 (2009).
- Beggs, J. M. & Plenz, D. Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**, 11167–11177 (2003).
- Mazzoni, A. et al. On the dynamics of the spontaneous activity in neuronal networks. *PLoS ONE* **2**, e439 (2007).
- Petermann, T. et al. Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. *Proc. Natl Acad. Sci. USA* **106**, 15921–15926 (2009).

7. Hahn, G. et al. Neuronal avalanches in spontaneous activity in vivo. *J. Neurophysiol.* **104**, 3312–3322 (2010).
8. Allegrini, P., Paradisi, P., Menicucci, D. & Gemignani, A. Fractal complexity in spontaneous EEG metastable-state transitions: new vistas on integrated neural dynamics. *Front. Physiol.* **1**, 128 (2010).
9. Tagliazucchi, E., Balenzuela, P., Fraiman, D. & Chialvo, D. R. Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Front. Physiol.* **3**, 15 (2012).
10. Beggs, J. M. The criticality hypothesis: how local cortical networks might optimize information processing. *Philos. Trans. R. Soc. A* **366**, 329–343 (2008).
11. Chialvo, D. R. Emergent complex neural dynamics. *Nat. Phys.* **6**, 744–750 (2010).
12. Hesse, J. & Gross, T. Self-organized criticality as a fundamental property of neural systems. *Front. Syst. Neurosci.* **8**, 166 (2014).
13. Priesemann, V. et al. Spike avalanches in vivo suggest a driven, slightly subcritical brain state. *Front. Syst. Neurosci.* **8**, 108 (2014).
14. Wiltng, J. & Priesemann, V. Inferring collective dynamical states from widely unobserved systems. *Nat. Commun.* **9**, 2325 (2018).
15. Bak, P. *How Nature Works: The Science of Self-Organised Criticality*. (Copernicus Press, New York, NY, 1996).
16. Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena*. (Oxford University Press, Oxford, 1987).
17. Stoop, R. & Gomez, F. Auditory power-law activation avalanches exhibit a fundamental computational ground state. *Phys. Rev. Lett.* **117**, 038102 (2016).
18. Beggs, J. M. & Timme, N. Being critical of criticality in the brain. *Front. Physiol.* **3**, 163 (2012).
19. Touboul, J. & Destexhe, A. Power-law statistics and universal scaling in the absence of criticality. *Phys. Rev. E* **95**, 012413 (2017).
20. Kanders, K., Lorimer, T. & Stoop, R. Avalanche and edge-of-chaos criticality do not necessarily co-occur in neural networks. *Chaos* **27**, 047408 (2017).
21. Shew, W. L. & Plenz, D. The functional benefits of criticality in the cortex. *Neuroscientist* **19**, 88–100 (2013).
22. Haldeman, C. & Beggs, J. M. Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Phys. Rev. Lett.* **94**, 058101 (2005).
23. Harris, T. E. *The Theory of Branching Processes*. (Dover Publications, New York, 1989).
24. Zapperi, S., Lauritsen, K. B. & Stanley, H. E. Self-organized branching processes: mean-field theory for avalanches. *Phys. Rev. Lett.* **75**, 4071–4074 (1995).
25. Tetzlaff, C., Okujeni, S., Egert, U., Wörgötter, F. & Butz, M. Self-organized criticality in developing neuronal networks. *PLoS Comput. Biol.* **6**, e1001013 (2010).
26. Shew, W. L. et al. Adaptation to sensory input tunes visual cortex to criticality. *Nat. Phys.* **11**, 659–664 (2015).
27. Ribeiro, T. L., Ribeiro, S., Belchior, H., Caixeta, F. & Copelli, M. Undersampled critical branching processes on small-world and random networks fail to reproduce the statistics of spike avalanches. *PLoS ONE* **9**, e94992 (2014).
28. Yaghoubi, M. et al. Neuronal avalanche dynamics indicates different universality classes in neuronal cultures. *Sci. Rep.* **8**, 3417 (2018).
29. Nerbonne, J. M. & Gurney, A. M. Development of excitable membrane properties in mammalian sympathetic neurons. *J. Neurosci.* **9**, 3272–3286 (1989).
30. Oswald, A. M. M. & Reyes, A. D. Maturation of intrinsic and synaptic properties of layer 2/3 pyramidal neurons in mouse auditory cortex. *J. Neurophysiol.* **99**, 2998–3008 (2008).
31. Zucker, R. S. & Regehr, W. G. Short-term synaptic plasticity. *Annu. Rev. Physiol.* **64**, 355–405 (2002).
32. Hó, N. & Destexhe, A. Synaptic background activity enhances the responsiveness of neocortical pyramidal neurons. *J. Neurophysiol.* **84**, 1488–1496 (2000).
33. Kim, R., Joo, S., Jung, H., Hong, N. & Nam, Y. Recent trends in microelectrode array technology for in vitro neural interface platform. *Biomed. Eng. Lett.* **4**, 129–141 (2014).
34. Wagenaar, D. A., Pine, J. & Potter, S. M. An extremely rich repertoire of bursting patterns during the development of cortical cultures. *BMC Neurosci.* **7**, 11 (2006).
35. Stewart, C. V. & Plenz, D. Homeostasis of neuronal avalanches during postnatal cortex development in vitro. *J. Neurosci. Methods* **169**, 405–416 (2008).
36. Azevedo, F. A. C. et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513**, 532–541 (2009).
37. Chung, W. S., Allen, N. J. & Eroglu, C. Astrocytes control synapse formation, function, and elimination. *CSH Perspect. Biol.* **7**, a020370 (2015).
38. Ichikawa, M., Muramoto, K., Kobayashi, K., Kawahara, M. & Kuroda, Y. Formation and maturation of synapses in primary cultures of rat cerebral cortical cells: an electron microscopic study. *Neurosci. Res.* **16**, 95–103 (1993).
39. Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703 (2009).
40. Deluca, A. & Corral, Á. Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophys.* **61**, 1351–1394 (2013).
41. Lorimer, T., Gomez, F. & Stoop, R. Two universal physical principles shape the power-law statistics of real-world networks. *Sci. Rep.* **5**, 12353 (2015).
42. Newman, M. E. J. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **46**, 323–351 (2005).
43. Friedman, N. et al. Universal critical dynamics in high resolution neuronal avalanche data. *Phys. Rev. Lett.* **108**, 208102 (2012).
44. Marshall, N. et al. Analysis of power laws, shape collapses, and neural complexity: new techniques and MATLAB support via the NCC toolbox. *Front. Physiol.* **7**, 250 (2016).
45. Sethna, J. P., Dahmen, K. A. & Myers, C. R. Crackling noise. *Nature* **410**, 242–250 (2001).
46. Fagerholm, E. D. et al. Cascades and cognitive state: focused attention incurs subcritical dynamics. *J. Neurosci.* **35**, 4626–4634 (2015).
47. Wechsler, D. & Stoop, R. in *Emergent Complexity from Nonlinearity, in Physics, Engineering and the Life Sciences* (eds Mantica, G. et al.), 105–126 (Springer, Cham, 2017).
48. Landis, F., Ott, T. & Stoop, R. Hebbian self-organizing integrate-and-fire networks for data clustering. *Neural Comput.* **22**, 273–288 (2010).
49. Gomez, F., Stoop, R. L. & Stoop, R. Universal dynamical properties preclude standard clustering in a large class of biochemical data. *Bioinformatics* **30**, 2486–2493 (2014).
50. Stoop, R., Kanders, K., Lorimer, T., Held, J. & Albert, C. Big data naturally rescaled. *Chaos Soliton. Fract.* **90**, 81–90 (2016).
51. Held, J., Lorimer, T., Albert, C. & Stoop, R. in *Emergent Complexity from Nonlinearity, in Physics, Engineering and the Life Sciences* (eds Mantica, G., Stoop, R., & Stramaglia, S.), 127–141 (Springer, Cham, 2017).
52. Lorimer, T., Held, J. & Stoop, R. Clustering: how much bias do we need? *Phil. Trans. R. Soc. A* **375**, 20160293 (2017).
53. Stoop, R., Bunimovich, L. A. & Steeb, W.-H. Generic origins of irregular spiking in neocortical networks. *Biol. Cybern.* **83**, 481–489 (2000).
54. Huffman, D. A. A method for the construction of minimum-redundancy codes. *Proc. IRE* **40**, 1098–1101 (1952).
55. Stoop, R. & Stoop, N. Natural computation measured as a reduction of complexity. *Chaos* **14**, 675–679 (2004).
56. van Kessenich, L. M., Luković, M., De Arcangelis, L. & Herrmann, H. J. Critical neural networks with short-and long-term plasticity. *Phys. Rev. E* **97**, 032312 (2018).
57. de Arcangelis, L. & Herrmann, H. J. Learning as a phenomenon occurring in a critical state. *Proc. Natl Acad. Sci. USA* **107**, 3977–3981 (2010).
58. Levina, A., Herrmann, J. M. & Geisel, T. Dynamical synapses causing self-organized criticality in neural networks. *Nat. Phys.* **3**, 857 (2007).
59. Uzunov, D. I. *Introduction to the Theory of Critical Phenomena*. (World Scientific Publishing Company, Singapore, 2010).
60. de Arcangelis, L., Perrone-Capano, C. & Herrmann, H. J. Self-organized criticality model for brain plasticity. *Phys. Rev. Lett.* **96**, 028107 (2006).
61. van Kessenich, L. M., de Arcangelis, L. & Herrmann, H. J. Synaptic plasticity and neuronal refractory time cause scaling behavior of neuronal avalanches. *Sci. Rep.* **6**, 32071 (2016).
62. Benayoun, M., Cowan, J. D., van Drongelen, W. & Wallace, E. Avalanches in a stochastic model of spiking neurons. *PLoS Comput. Biol.* **6**, 1000846 (2010).
63. Orlandi, J. G., Soriano, J., Alvarez-Lacalle, E., Teller, S. & Casademunt, J. Noise focusing and the emergence of coherent activity in neuronal cultures. *Nat. Phys.* **9**, 582–590 (2013).
64. Andreae, L. C. & Burrone, J. The role of neuronal activity and transmitter release on synapse formation. *Curr. Opin. Neurobiol.* **27**, 47–52 (2014).
65. Spitzer, N. C., Kingston, P. A., Manning, T. J. Jr & Conklin, M. W. Outside and in: development of neuronal excitability. *Curr. Opin. Neurobiol.* **12**, 315–323 (2002).
66. Pratt, K. G. & Aizenman, C. D. Homeostatic regulation of intrinsic excitability and synaptic transmission in a developing visual circuit. *J. Neurosci.* **27**, 8268–8277 (2007).
67. Jackman, S. L. & Regehr, W. G. The mechanisms and functions of synaptic facilitation. *Neuron* **94**, 447–464 (2017).
68. Dayan, P. & Abbott, L. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. (MIT Press, Cambridge, MA, 2001).

69. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap*. (Chapman and Hall, New York, 1993).
70. Maaten, L. V. D. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).

Acknowledgements

The research was supported by a Swiss-Korea collaboration grant (IZKS2_162190), an internal grant of ETH Zürich (ETH-37 152) and by a National Research Foundation of Korea grant (NRF-2015K1A3A1A14021018).

Author contributions

R.S and Y.N. conceived the experiments, H.L. and N.H. conducted the experiments, K.K. and R.S. conceived the analysis, developed the model and wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42005-019-0276-8>.

Correspondence and requests for materials should be addressed to K.K. or R.S.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020