

RESEARCH ARTICLE

Signatures of criticality arise from random subsampling in simple population models

Marcel Nonnenmacher^{1,2,3*}, Christian Behrens^{3,4,5}, Philipp Berens^{3,4,5}, Matthias Bethge^{2,3,4,6}, Jakob H. Macke^{1,2,3*}

1 Research Center caesar, an associate of the Max Planck Society, Bonn, Germany, **2** Max Planck Institute for Biological Cybernetics, Tübingen, Germany, **3** Bernstein Center for Computational Neuroscience, Tübingen, Germany, **4** Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany, **5** Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany, **6** Institute of Theoretical Physics, University of Tübingen, Tübingen, Germany

* marcel.nonnenmacher@caesar.de (MN); jakob.macke@caesar.de (JHM)



OPEN ACCESS

Citation: Nonnenmacher M, Behrens C, Berens P, Bethge M, Macke JH (2017) Signatures of criticality arise from random subsampling in simple population models. *PLoS Comput Biol* 13(10): e1005718. <https://doi.org/10.1371/journal.pcbi.1005718>

Editor: Peter E. Latham, UCL, UNITED KINGDOM

Received: February 17, 2017

Accepted: August 1, 2017

Published: October 3, 2017

Copyright: © 2017 Nonnenmacher et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data used for our study was simulated, with some summary statistics extracted from publications cited within our manuscript. We uploaded our simulated data, the code used to generate the data, as well as code for data analysis and generation of the main figures to: https://github.com/mackelab/critical_retina

Funding: Work was funded by the German Federal Ministry of Education and 568 Research (BMBF; FKZ: 01GQ1002, Bernstein Center Tübingen, FKZ 01GQ1601 to PB), the 569 German Research Foundation (BE 5601/1-1 to PB) the Max Planck

Abstract

The rise of large-scale recordings of neuronal activity has fueled the hope to gain new insights into the collective activity of neural ensembles. How can one link the statistics of neural population activity to underlying principles and theories? One attempt to interpret such data builds upon analogies to the behaviour of collective systems in statistical physics. Divergence of the specific heat—a measure of population statistics derived from thermodynamics—has been used to suggest that neural populations are optimized to operate at a “critical point”. However, these findings have been challenged by theoretical studies which have shown that common inputs can lead to diverging specific heat. Here, we connect “signatures of criticality”, and in particular the divergence of specific heat, back to statistics of neural population activity commonly studied in neural coding: firing rates and pairwise correlations. We show that the specific heat diverges whenever the average correlation strength does not depend on population size. This is necessarily true when data with correlations is randomly subsampled during the analysis process, irrespective of the detailed structure or origin of correlations. We also show how the characteristic shape of specific heat capacity curves depends on firing rates and correlations, using both analytically tractable models and numerical simulations of a canonical feed-forward population model. To analyze these simulations, we develop efficient methods for characterizing large-scale neural population activity with maximum entropy models. We find that, consistent with experimental findings, increases in firing rates and correlation directly lead to more pronounced signatures. Thus, previous reports of thermodynamical criticality in neural populations based on the analysis of specific heat can be explained by average firing rates and correlations, and are not indicative of an optimized coding strategy. We conclude that a reliable interpretation of statistical tests for theories of neural coding is possible only in reference to relevant ground-truth models.

Society and the caesar foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Understanding how populations of neurons collectively encode sensory information is one of the central goals of computational neuroscience. In physics, systems are often characterized by identifying and describing critical points (e.g. the transition between two states of matter). The success of this approach has inspired a series of studies to search for analogous phenomena in nervous systems, and has led to the hypothesis that these might be optimized to be poised at ‘thermodynamic critical points’. However, translating concepts from thermodynamics to neural data analysis has been a challenging endeavour. We here study the data analysis approaches that have been used to provide evidence for criticality in the brain. We find that observing signatures of criticality is closely linked to observing activity correlations between neurons— a ubiquitous phenomenon in neural data. Our study questions the experimental evidence that neural systems are optimised to exhibit thermodynamic critical behaviour. Finally, we provide practical, open-source tools for analyzing large-scale measurements of neural population activity using maximum entropy models.

Introduction

Recent advances in neural recording technology [1, 2] and computational tools for describing neural population activity [3] make it possible to empirically examine the statistics of large neural populations and search for principles underlying their collective dynamics [4]. One hypothesis that has emerged from this approach is the idea that neural populations might be poised at a thermodynamic critical point [5, 6, 7], and that this might have consequences for how neural populations process sensory information [7, 8]. As similar observations have been made in other biological systems [9, 10, 11], it has been suggested that this might reflect a more general organising principle [12]. Critical phenomena play a central role in physics: Phase transitions mark a special point in which media qualitatively change their properties by transitioning from one state of matter into another (e.g. liquid to gaseous at boiling point, ferro-magnetic and paramagnetic phases, or the emergence of super-conductivity). As such, the behaviour of a system at critical points is informative about its intrinsic properties. Moreover, critical points are ‘special’ in the sense that they classically only occupy a small portion of the parameter space. Thus, observing that a system is constantly poised at a critical point would be surprising, and would hint at an underlying organizing mechanism that keeps the system at this point. Given the fundamental importance of critical phenomena in physics, and their success in revealing the laws that determine the behaviour of physical systems, the hypothesis that these approaches might also shed lights on principles underlying neural coding is intriguing.

Evidence in favour of this hypothesis has been put forward by a series of studies which measured neural activity from large populations of retinal ganglion cells and reported that their statistics resemble those of physical systems at a critical point [7, 8]. To this end, Tkačik and colleagues developed a data analysis framework to search for signatures of criticality in experimentally obtained measurements. Using large-scale multielectrode array recordings [2] and maximum entropy models [13, 14, 15, 16, 17, 12, 3, 18], it was observed that the normalized variance of log-probabilities diverges as a function of population size. Importantly, this quantity is mathematically equivalent to the specific heat capacity, an important characteristic which diverges at critical points. In addition, when an artificial ‘temperature’ parameter was introduced, specific heat appeared to be maximal for the statistics of the observed data, rather

than for statistics which have been perturbed by changing the temperature parameter. These properties of retinal populations resemble the behaviour of physical systems at critical points. It has been hypothesised [12, 7] that the system needs to be optimized to keep itself at a critical point, for example through adaptation to stimulus statistics [19, 20, 21] or alternative mechanisms of self-organization [22, 23, 24].

A competing hypothesis states that instead generic mechanisms are sufficient to give rise to activity data with divergent specific heat, and that the presence of signatures of criticality does not provide evidence for retinal circuits being poised at a special state that is advantageous for coding. A series of theoretical studies [25, 26, 27, 28] has shown that common input (i.e., the presence of latent variables) can account for signatures of criticality: In particular, Schwab et al [26] and Aitchison et al [27, 28] showed that Zipf scaling (an alternative characterization of criticality) and the divergence of the specific heat are closely related, and that in high-dimensional models with a low-dimensional latent variable, the specific heat diverges with system size under a wide range of circumstances [27, 28]. Similarly, it has been shown empirically that a purely feedforward model can capture Zipf-like scaling in recordings from the salamander retina [29].

Interpreting findings of thermodynamic criticality for neural populations, identifying their mechanistic underpinnings, and clarifying their relationship with alternative theories, has been fraught with difficulty. We hypothesize that this difficulty stems from a subtle but crucial difference between how the scaling behaviour of system properties is studied in thermodynamics and in practical neural data analysis: Most theoretical approaches study how system properties scale as the size of the system, n , is varied. In contrast, in practical neural data analysis, different “ n ” do not correspond to different system sizes, but are obtained by subsampling neural populations from a large recording (which is itself a subsample of the underlying system). How does this sampling process affect estimates of whether the system is at a critical point? A second difficulty in interpreting these studies stems from the fact that they are based global statistical measures whose relationship with simple statistics such as firing rates and correlations— which are commonly used and have been extensively studied in neural coding [30, 31]—is unclear. We here focus on one statistic that has been used as evidence of critical behaviour, namely the dependence of specific heat on population size and temperature. We study how it depends on neural firing rates and correlations, as well as on how this data is subsampled during data analysis:

First, we show explicitly that signatures of criticality, can be reproduced in canonical feed-forward models of neural population activity, as predicted by previous studies [25, 26, 28]. These studies did not have tools for studying population statistics in large simulations, and they were therefore limited to studying small ($n \leq 40$) systems— for these small system sizes, it is difficult to make statements about the peak in the specific heat and its scaling with population size. In particular, the dominant peak near unit temperature only emerges for much larger systems. We overcome this difficulty by providing improved algorithms for efficiently fitting maximum entropy models to large neural populations (available at <https://github.com/mackelab/CorBinian>), and use them to apply the analyses proposed by previous studies [7] to data simulated from a simple, feedforward encoding model of retinal processing [32, 33, 34, 35].

Second, previous theoretical studies [26, 27, 28] treated only the limiting behavior of the specific heat at unit temperature, and did not investigate its dependence on firing rates and correlations. We here relate the characteristic shape of specific heat curves (i.e. the dependence of specific heat on temperature) to neural correlations and firing rates. The emergence of peak specific heat at the ‘inherent’ temperature $T = 1$ has given rise to the idea that correlations in the observed system are ‘special’, i.e. that systems with stronger or weaker correlations would not exhibit them [7]. We use an analytically tractable model of the analysis process to show

that this is not the case— the more strongly correlated the population is, the more pronounced signatures of criticality will be. This analysis also shows that a ‘low-temperature’ regime (as reported by [18]) will be found whenever firing rates are sufficiently low.

Third, we analyze the structure of correlations which are sufficient to induce signatures of criticality, and find that it is sufficient if the average correlation is independent of population size. Such ‘criticality-inducing’ correlations can arise both from neural mechanisms such as common input or dense connectivity. Importantly, we show that they can also arise as a consequence of data analysis: Uniformly subsampling a recording with any non-zero correlations to construct subpopulations yields criticality-inducing correlations.

In summary, we show that statements about signatures of criticality derived from thermodynamics can be reduced to statements about firing rates and correlations, and that correlation structures which give rise to these signatures are ubiquitous in neural populations.

Results

Signatures of criticality arise in a simple model of retinal ganglion cell activity

A hallmark of criticality is that the specific heat capacity of the model diverges when the temperature reaches the critical temperature [5]. Tkačik et al. [7] developed an approach for translating this concept to neural data analysis (see Fig 1):. In this analysis, neural populations of different size n are generated from the full recording (of size N) by random subsampling. The statistics of activity for each population of size n are characterized using a maximum entropy

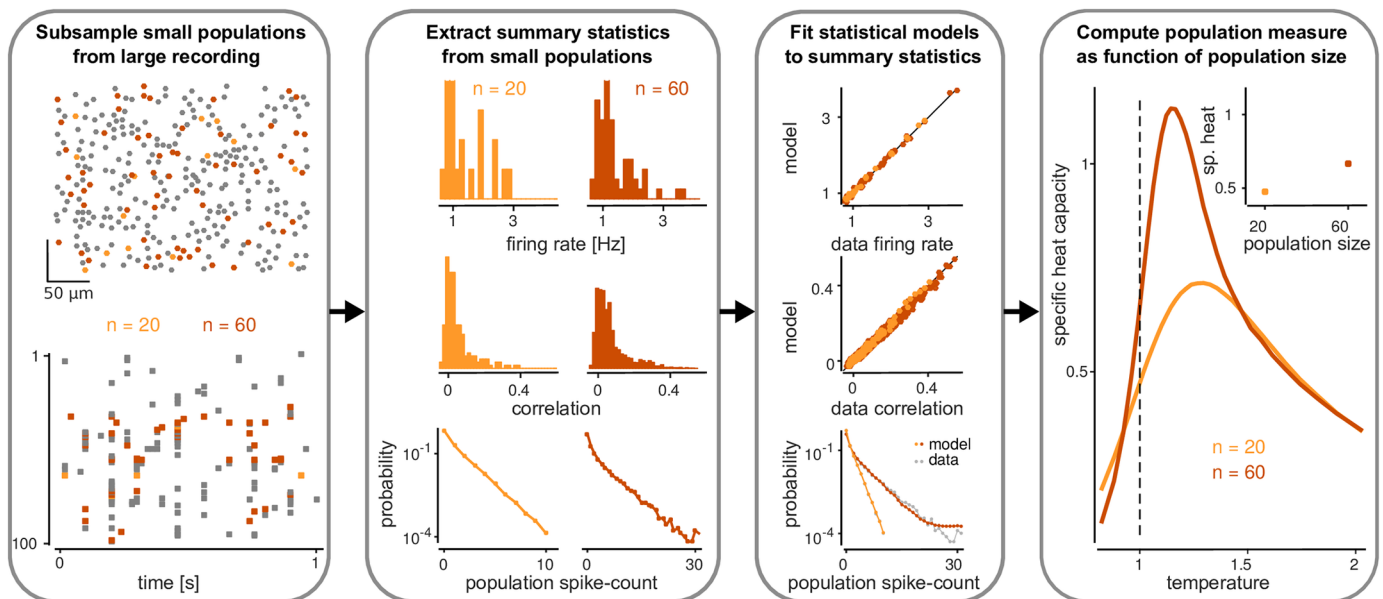


Fig 1. How can one relate theories of thermodynamic criticality to the statistics of neural data? In physical systems, the divergence of specific heat with system size can be interpreted as the system being at a critical point. We here study an analysis approach that has been proposed in order to search for similar signatures of criticality in the statistics of neural population activity. In this approach, different populations are subsampled from a large recording and summary statistics are extracted for each subpopulation (e.g. firing rates, correlations and population spike count statistics). Subsequently, maximum entropy models are fit to these data which assign a probability to each possible spike-pattern. Exploiting the mathematical relationship between the log-variance of probabilities (in statistics) and the specific heat (in thermodynamics) then allows one to compute and study the behaviour of the specific heat with population size. The goal of this study is to determine under which conditions (i.e., for which firing rates and correlations) such an analysis would report that the system is critical. To this end, we apply this approach to a simulation of neural population activity and analytically tractable models.

<https://doi.org/10.1371/journal.pcbi.1005718.g001>

model fit to population activity [13, 14, 16, 17, 3]. Finally, the maximum entropy models are perturbed by introducing a temperature parameter, and specific heat is computed for each population size n and temperature T from the (perturbed) maximum entropy model fit. Divergence of specific heat with population size n , and a peak of the specific heat near unit temperature $T = 1$ (the ‘temperature’ of the original data) are interpreted as indication for the system being at a critical point [7].

We wanted to verify that this phenomenon could be captured in feedforward models of retinal processing. We wanted to directly demonstrate that canonical mechanisms of retinal processing—such overlapping centre-surround receptive fields, spiking nonlinearities, shared Gaussian noise—are sufficient for the signatures of criticality to arise. We first created a simple phenomenological model of retinal ganglion cell (RGC) activity based on linear-nonlinear neurons [32, 33, 35]. In this model (Fig 2a), we assumed retinal ganglion cells to have centre-surround receptive fields [36, 35] with linear spatial integration [37], sigmoid nonlinearities and stochastic binary spikes: in each time bin of size 20ms, each neuron i either emitted a spike ($x_i = 1$) or not ($x_i = 0$). We used a sequence of natural images as stimuli. In addition to the feedforward drive by the stimulus, nearby neurons received shared Gaussian noise, mimicking common input from bipolar cells [30]. Thus, cross-neural correlations in the model arise from correlations in the stimulus, receptive-field overlap and shared noise, but not from lateral connections between RGCs. As we will explain below, only the strength of correlations, but not their mechanistic origin or dependence on stimuli, is relevant for determining the specific heat. Parameters of the model were chosen to approximate the statistics of receptive-field centre locations of RGCs, as well as histograms of firing rates, pairwise correlation-coefficients and population spike-counts (Fig 2b).

We subsampled populations of different sizes $20 \leq n \leq 120$ by uniformly sampling cells from our simulated recording of total size $N = 316$ neurons. For each population we fit a ‘K-pairwise’ maximum entropy model [3]. This model assigns a probability $P(\mathbf{x})$ to each spike-pattern \mathbf{x} . It is an extension of pairwise maximum entropy models (i.e. Ising models) [13, 14] which reproduce the firing rates and pairwise covariances, and has additional terms to capture population spike-counts [3] (see Materials for details of model specification and parameterisation). As we needed to efficiently fit this model [38, 39] to multiple simulated data sets, we developed an improved fitting algorithm (see section 1 in S1 Supporting Information) based on maximum-likelihood techniques using Markov chain Monte Carlo (MCMC), building on work by [15]. In particular, we made the most computationally expensive component of the algorithm, the estimation of pairwise covariances via MCMC sampling, more efficient by using a ‘pairwise’ Gibbs-sampling scheme with Rao-Blackwellisation [40] (see section 1.1 in S1 Supporting Information). Most Gibbs-sampling approaches for maximum entropy models [15] update one neuron i at a time by re-sampling its state from the conditional distribution, given the state of the other $n - 1$ neurons in the population. We here in each iteration update a randomly chosen pair (i, j) simultaneously, given the state of the other $n - 2$ neurons. While each pairwise sample is more expensive to compute, this approach has the advantage of yielding a direct estimate of the (conditional) probability of i and j being active simultaneously. From these conditional probabilities, one can estimate pairwise covariances more efficiently than is possible through averaging samples, a process which is known as Rao-Blackwellization. Here, Rao-Blackwellization resulted in a reduction of the number of samples (and computation time) needed for achieving low-variance estimates of the covariances by a factor of approximately 3 (Fig 2c, Fig. A in S1 Supporting Information). After parameter fitting, the model reproduced the statistics of the simulated data (Fig 2d, Fig. B in S1 Supporting Information).

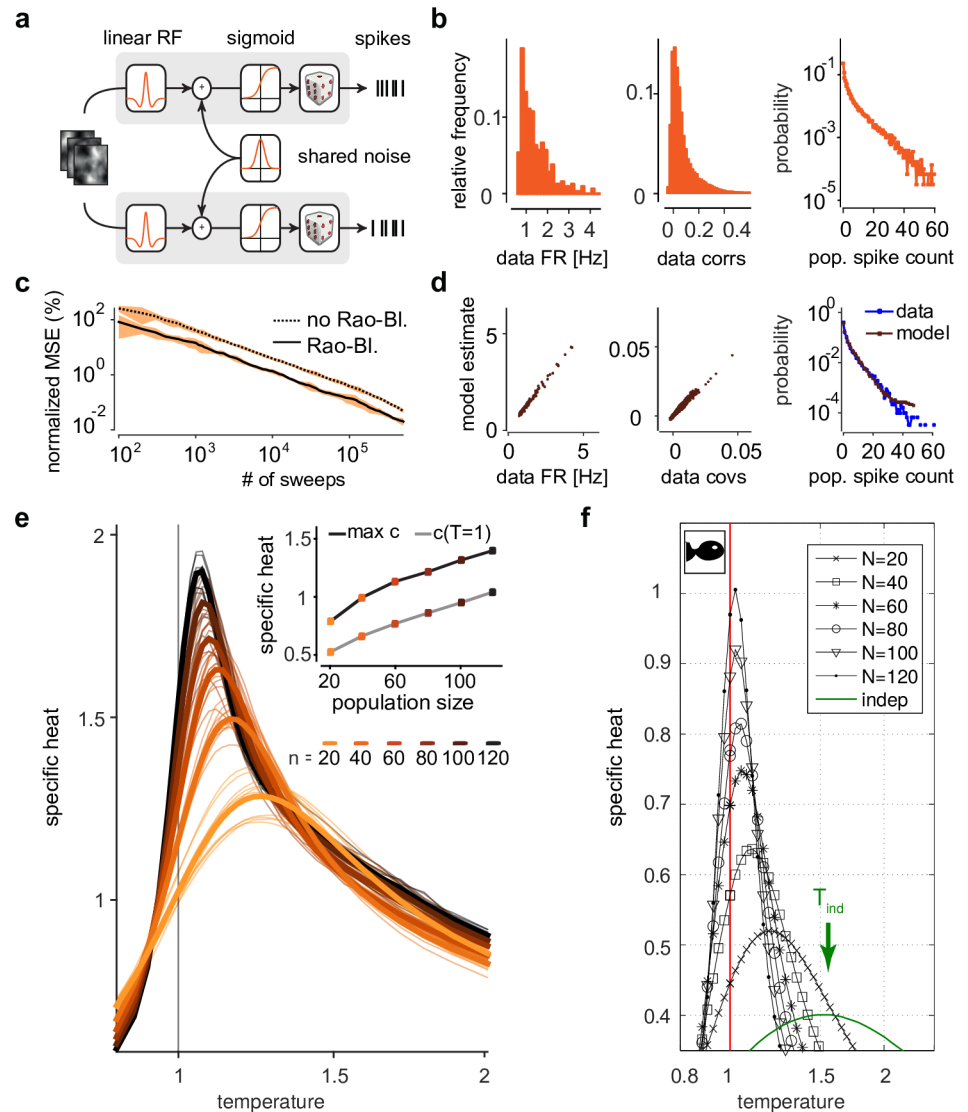


Fig 2. Signatures of criticality in a simulation of retinal ganglion cell activity. **a)** Simulation schematic: Neurons have linear stimulus selectivity with centre-surround receptive fields and correlated Gaussian noise. **b)** Statistics of simulated population activity. Histograms of firing rates (left), correlation coefficients (centre) and frequency of population spike-counts (right). **c)** Estimation-error (normalised mean square error) in pairwise covariances as function of sample size, averaged across 10 populations of size $n = 100$. Rao-Blackwellization reduces the number of samples needed for a given level of accuracy by a factor ≈ 3 . **d)** Quality of fit: Population models (here $n = 100$, example population) capture the mean firing rates (left), covariances (centre) and spike-counts (right). **e)** Divergence of specific heat: Average and individual traces for 10 randomly sampled populations for each of 6 different population sizes, exhibiting divergence of specific heat and peak in heat near unit temperature. Inset: Specific heat at unit temperature and at peak vs. population size. **f)** Specific heat for different temperatures and subsampled population sizes (here denoted by capital letter N) in recordings of salamander retinal ganglion cells responding to naturalistic stimuli, reproduced from [7].

<https://doi.org/10.1371/journal.pcbi.1005718.g002>

Following [7], we then introduced a temperature parameter which rescales the probabilities of the model,

$$P_T(\mathbf{x}) \propto P(\mathbf{x})^{1/T}, \quad (1)$$

where temperature $T = 1$ corresponds to the statistics of the empirical data. By changing T to other parameter values one can perturb the statistics of the system [41]: Increasing temperature leads to models with higher firing rates and weaker correlations (Fig. C in [S1 Supporting Information](#)), with $P_T(\mathbf{x})$ approaching the uniform distribution for large T . If the temperature is decreased towards zero, $P_T(\mathbf{x})$ has most of its probability mass over the most probable spike patterns. We compute the specific heat of a population directly from the probabilistic model fit to data [7], using

$$c(T) = \frac{1}{n} \text{Var}[\log P_T(X|\lambda)], \quad (2)$$

i.e. the variance of the log-probabilities of the model with parameters λ , normalised by n . While specific heat is typically motivated by thermodynamics, in this context it corresponds to a global statistical measure which provides a compact mathematical description of the collective statistical dynamics of the system. Just like the entropy corresponds to the (negative) average log-probability across all population states, the specific heat corresponds to the (normalized) variance of log-probabilities. Thus, specific heat is minimal for data in which all patterns \mathbf{x} are equally probable, and big for data in which pattern-probabilities span a large range. We used MCMC-sampling to approximate the variance across all probabilities, and used this approach to calculate, for each population of size n , the specific heat as a function of temperature (Fig. D in [S1 Supporting Information](#)).

We found that the temperature curves obtained from the simulated data qualitatively reproduce the critical features of those that had been observed for large-scale recordings in the salamander [7] and rat [8] retina: The peak of the curves diverges as the population size n is increased, and moves closer to unit temperature for increasing n (Fig 2e). Consistent with experimental findings [42, 7, 8] (Fig 2f) and [28], we found that specific heat diverged linearly with population size. Finally, and also consistent with experimental studies, the peak specific heat is achieved for $T > 1$, which is what has been interpreted as a ‘low-temperature’ state [18]. These results confirm that signatures of criticality arise in a simple feedforward LN cascade model based on generic properties of retinal ganglion cells, and do not require finely tuned parameters or sophisticated circuitry.

A tractable mathematical model of the analysis process explains specific-heat curves and low-temperature states

In the phenomenological population model above, we observed that specific heat grew linearly with population size, as it did in previous studies built on experimental data [42, 7, 8, 18]. Different ‘populations’ in these analyses are obtained by subsampling different populations from a large experimental recording, and that the parameters of each of these models are independently fit to each such population. How does this analysis process effect the rate of divergences of the specific heat, and the qualitative shape of specific heat curves? To answer these questions, we build a simple mathematical description of the analysis process: In the original papers, populations of different sizes are obtained by randomly subsampling a large recording (which is itself a sub-sample of the underlying circuit). As the simplest possible description of this sampling process, we assume that there is an underlying, infinitely large neural population, and that each population of size n is a random subsample. We assume that the underlying population is homogeneous, i.e. that all neurons have the same mean firing rate and pairwise correlations. As a consequence, K -pairwise maximum entropy models are fully specified by the distribution of population spike-count $K = \sum_i x_i$ [25, 43, 44, 45] for each population of size n .

We refer to models with this property as ‘flat models’ ([46] calls them ‘reduced’ maximum entropy models).

We introduce a new parametrised flat model in which the spike-count distribution is given by the beta-binomial distribution $P(K|\alpha, \beta, n)$, reducing the number of free parameters from n to 2. The beta-binomial model is a straightforward extension of an independent (i.e. binomial) population model: At each time-point, a new firing probability p is drawn from a beta-distribution with parameters α and β , and neurons then spike independently with probability p . Fluctuations in the latent variable p are shared across the population and lead to correlations in neural activity. Therefore, this model is a particular instance of a latent variable model. Signatures of criticality in latent variable models have been studied previously [26, 27, 28]. Our analytically-tractable model provides an explicit construction of how subsampling a large population determines the dependence of specific heat on population size.

Our beta-binomial model provided a good fit to the population spike-count distributions of the simulated data (Fig 3a) across different population sizes n (Fig 3b). Importantly, the best-fitting parameters α and β did not vary systematically across population sizes, and converged to values of $\alpha = 0.38$ and $\beta = 12.35$ (Fig. E in S1 Supporting Informationa), corresponding to a probability of spiking of $\mu = 0.03$ in each bin (i.e. each neuron has an average

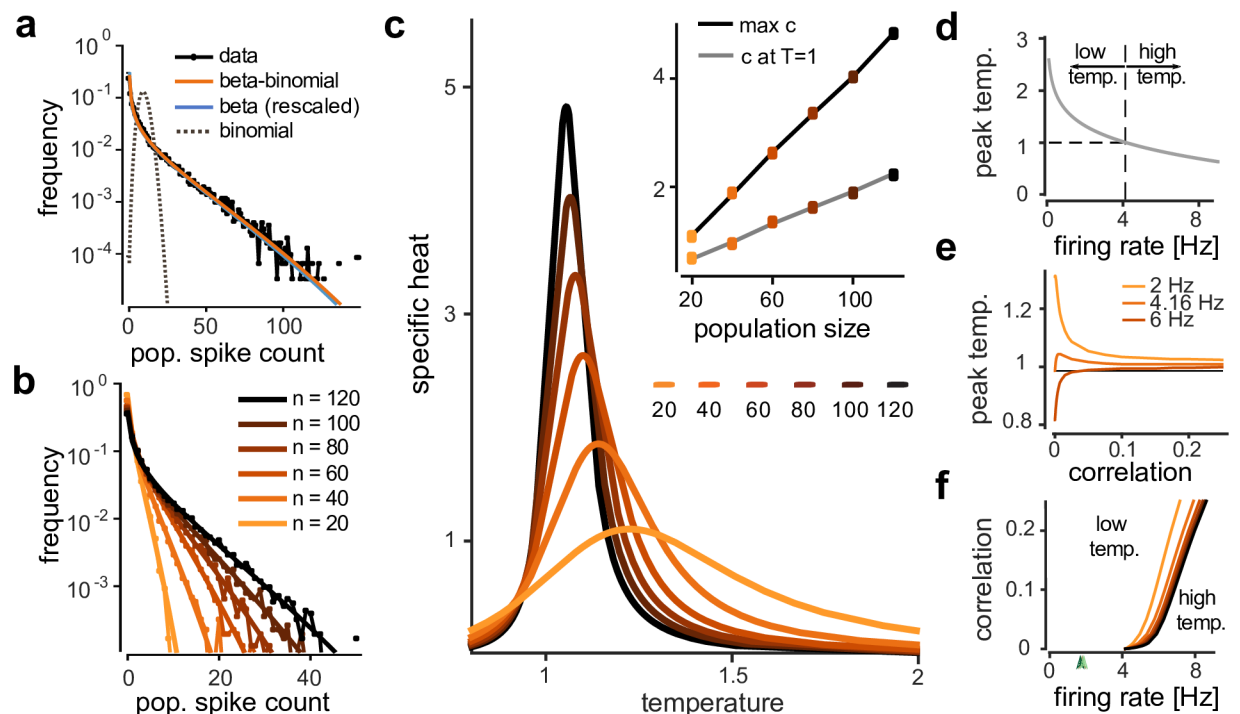


Fig 3. Signatures of criticality and low-temperature states in a mathematically tractable model. **a)** Population spike-count distribution in RGC simulation, and approximation by models. Only the beta-binomial population model fits simulated data accurately, and for the full recording ($N = 316$) closely matches the shape of a beta distribution. **b)** Beta-binomial model fits for different population sizes. **c)** Specific heat traces for beta-binomial model, exhibiting signatures of criticality. Average and individual traces for 30 randomly sampled populations for each of 6 different population sizes. Inset: Specific heat at unit temperature and at peak vs. population size. **d)** Location of peak specific heat for independent model as function of firing rate. For $\mu/\Delta = 4.16\text{Hz}$ (assuming $\Delta = 20\text{ms}$ bins), the peak is above unit temperature, a ‘low-temperature phase’. **e)** Location of peak specific heat as function of correlation, for $n = 100$ and three different firing rates. Peaks cross $T = 1$ only for firing rates $\geq 4.16\text{Hz}$. **f)** ‘Low’ and ‘high’ temperature phases for beta-binomial model as function of firing rate and correlation strength and for population sizes ($n = 20$ to $n = 120$, colors as in **b,c**). Increasing correlations and population size expand the low-temperature regime beyond 4.16Hz . Data sets from previous studies had average firing rates well within low-temperature regime (arrows, colors as in Fig. E in S1 Supporting Information).

<https://doi.org/10.1371/journal.pcbi.1005718.g003>

firing rate of $\mu/\Delta = 1.5$ Hz) and average pairwise correlations of $\rho = 0.073$. The beta-binomial model also provided good fits to published population spike-count distributions [43, 45, 8], as well as to those of retinal ganglion cell activity under different stimulus conditions in [18] (Fig. E in [S1 Supporting Information](#)). When we applied this flat model to populations subsampled from the RGC simulation, we could qualitatively reproduce the specific heat curves of the K-pairwise model (see also Fig. F in [S1 Supporting Information](#)). In particular, we found a linearly diverging peak that moved closer to $T = 1$ as the population size was increased (Fig 3c). Thus, linear divergence of specific heat is qualitatively captured by this model of how different populations are obtained by subsampling a large population.

One of the difficulties of interpreting the scaling behaviour of maximum entropy models fit to neural data is the fact that the construction of the limit in n differs from those studied in statistical physics: In statistical physics, different ‘ n ’ typically correspond to systems of different total size, and the parameters are scaled as a deterministic function of n (e.g. drawn from a Gaussian with variance proportional to $1/n$ in spin-glasses [47, 48]). In studies using maximum entropy models for neural data analysis, populations of different n are obtained by randomly subsampling a fixed large recording, and the parameters are fit to each subpopulation individually. Thus, there is no analytical relationship between population size and parameter values in this approach. With our model of the analysis process based on flat models, it is possible to analytically characterise the behaviour of the specific heat for large population sizes for this sampling process [25, 44]. Using this approach, one can show (section 2.3 in [S1 Supporting Information](#) and [25] for details) that for virtually all flat models, the specific heat diverges linearly at unit temperature, but not for any other temperature $T > 1$ or $T < 1$ (section 2.4 in [S1 Supporting Information](#)). As a consequence, the peak must move to $T = 1$ as n is increased. Hence, almost any flat model analysed with the methods developed by [7] will exhibit signatures of criticality. In particular, these results hold also for models which are more weakly or more strongly correlated than real neural populations, and even for models with unrealistic population spike-count distributions (see Fig. G in [S1 Supporting Information](#) for an illustration). There are only two exceptions: The first one is a model in which all neurons are independent (i.e. a binomial population model), and the second one is a flat pairwise maximum entropy model—indeed, this is the only flat model with non-vanishing correlations for which the specific heat does not have its peak at unit temperature (see [25] for an illustration for the flat pairwise maximum entropy model).

Finally, it has been observed that the peak of the specific heat curve is consistently ‘to the right’ of $T = 1$, which was interpreted as the neural population activity in the retina being in a ‘low-temperature state’ [18]. Our analysis based on the flat model gives insights into this phenomenon: For correlation $\rho = 0$, the position of the peak can be calculated in closed form (Fig 3d). We observe that the peak will be at temperatures > 1 whenever the spike probability is smaller than $\mu^* = 0.0832$, which corresponds to a firing rate of $\mu^*/\Delta = 4.16$ Hz at a bin size of $\Delta = 20$ ms. Thus, in our model, the ‘temperature-state’ of a population can be reduced to a statement about the firing rate relative to the bin size used for analysis: For $\rho > 0$ (Fig 3e) and for larger population sizes n , the firing rate at which the transition occurs are shifted to slightly higher firing rates, i.e. the ‘low-temperature’ regime is even bigger, and e.g. extends to firing rates up to 8.63 Hz for average correlations of $\rho = 0.25$ and population size $n = 120$ (Fig 3f). While this dependence may be more complicated for full correlation structures, our analysis again connects global population measures from statistical mechanics to basic, directly measurable statistics of neural data: ‘being in a low-temperature state’ is a statement about the firing rates in the population being low.

Strong neural correlations lead to fast divergence of specific heat

The rate at which the specific heat diverges provides a mean of quantifying the ‘strength’ of criticality. What is the relationship between correlations in a neural population and the rate of divergence? To study how the specific heat rate $\tilde{c} = c(T = 1)/n$ depends on the strength of correlations, we used a beta-binomial model to generate simulated data with firing rate $\mu/\Delta = 1.5\text{Hz}$ (i.e. each neuron has a probability of spiking of $\mu = 0.03$ per bin), and different pairwise correlation coefficient ρ ranging from $\rho = 0.01$ to $\rho = 0.25$ (Fig 4a). The heat curves had the same shape as in the analyses above, with a peak that increases and moves to unit temperature (Fig 4b). We found that the specific heat rates \tilde{c} increased strictly monotonically with ρ (Fig 4b and 4c). For the beta-binomial model, the large- n value of \tilde{c} can be calculated analytically (section 3.2 in S1 Supporting Information for details) as a function of the parameters α and β ,

$$\tilde{c} = \frac{\alpha(\alpha + 1)\psi_1(\alpha + 1) + \beta(\beta + 1)\psi_1(\beta + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} + \frac{\alpha\beta(\psi_0(\alpha + 1) - \psi_0(\beta + 1))^2}{(\alpha + \beta)^2(\alpha + \beta + 1)} - \psi_1(\alpha + \beta + 1), \tag{3}$$

where ψ_0, ψ_1 denote the di- and trigamma function, respectively. This analytical evaluation of \tilde{c} (valid for large n) was in good agreement with numerical simulations (Fig 4c left). In the case of weak correlations ρ , eq 3 can be simplified: In this case, the specific heat rate is proportional to the strength of correlations (section 3.1 in S1 Supporting Information for details), i.e.

$$\tilde{c} \approx \rho\mu(1 - \mu) \left(\log \left(\frac{1 - \mu}{\mu} \right) \right)^2, \tag{4}$$

and also increases strongly with firing rate for small μ (Fig. H in S1 Supporting Information). This expression can also be derived from the Gaussian model in [8] equation (4), by inserting the expected values of the mean and variance of the population spike-count under random subsampling. The monotonic relationship between correlations and specific heat is also consistent with the derivation in [27] for latent-variable models: inspection of equation (65) in [27] shows that the specific heat is related to a sum of conditional entropies— for binary random variables, these entropies are monotonically related to covariances, which effectively shows that, in their model, specific heat also increases with correlations.

We found that the relationship between the strength of correlations and the ‘strength’ of criticality (i.e. the divergence rate of specific heat) also held in simulations of feedforward models of retinal population activity. In the original study [7], specific heat was computed from K-pairwise model fits to RGC activity resulting from three different kind of stimuli: random checkerboard stimuli (which do not have long-range spatial correlations, although stimulus-driven cross-neural correlations can arise from receptive field overlap), natural stimuli, which exhibit strong spatial correlations, and full-field flicker (which constitutes an extreme case of spatial correlations since all pixels in the display are identical). It was found that specific heat diverges in all three conditions (consistent with a more recent study [18]), and interpreted this as evidence that signatures of criticality are not ‘inherited from the stimulus’ [7]. When we simulated responses to different stimuli we found the divergence rates of the specific heat to follow the pattern of induced correlation strength, consistent with the monotonic relationship between correlation strength and specific heat growth rate shown above for the flat models (Fig 4d): For populations size $n = 100$, checkerboard/natural/full-field flicker stimulation lead to average correlation strengths of $\rho = 0.033/0.075/0.341$, respectively, and to specific heat growth rates of $\tilde{c} = 0.0029/0.0046/0.0104$.

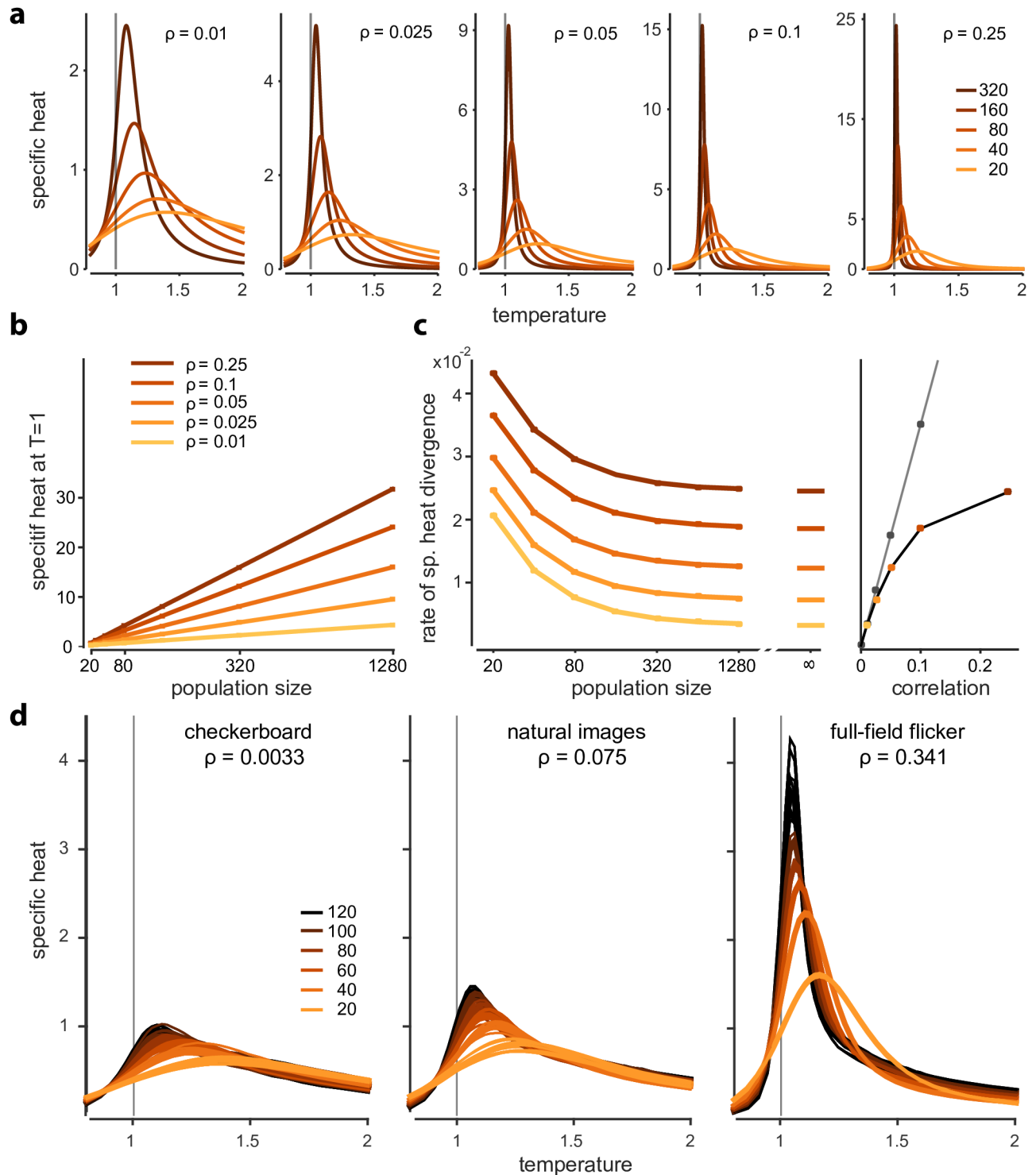


Fig 4. Relationship between correlations and criticality. **a**) Specific heat traces for beta-binomial model, different correlation strengths and population sizes. Heat traces are qualitatively similar, but differ markedly quantitatively (see y-axes). **b**) Specific heat diverges linearly, and the slope depends on the strength of correlations. **c**) Divergence rate of specific heat for beta-binomial model as a function of correlation strength (left). Rightmost point (at infinity) corresponds to analytical prediction of large- n behaviour. Divergence rates are strictly increasing with correlation strength (right) which is captured by a weak-correlation approximation (dashed line). **d**) Specific heat increases with correlation in the K-pairwise maximum entropy model: average and individual traces for 10 randomly subsampled populations for 6 different population sizes. Left to right: checkerboard, natural images and full-field flicker stimuli presented to the population. Correlation strengths denote mean correlation coefficient in each population.

<https://doi.org/10.1371/journal.pcbi.1005718.g004>

Tkačik et al. had found the lowest peak in divergence rate for checkerboard (max $c \approx 0.54$), higher peak-divergence rates for natural movies (max $c \approx 0.92$) and the highest peak for full-field flicker (max $c \approx 2.4$, all results for $n = 100$). Thus, the ordering of the peak values of specific heat in their study is consistent with our results. However, when comparing the values at $T = 1$, they found a slightly higher divergence rate for natural movies ($\tilde{c} \approx 0.005$) than for full-field flicker ($\tilde{c} \approx 0.004$). This mismatch could result from adaptation or temporal dynamics of the stimulus affecting firing rates or correlations in their data [20], or from our simulations not precisely matching the statistics of their experimental data.

These statements also qualitatively hold in a modified temperature analysis [7] in which firing rates are kept constant (at the firing rates of $T = 1$) when temperature is varied (section 3.4 in [S1 Supporting Information](#) and in Fig. I in [S1 Supporting Information](#)). We conclude that the experimental evidence—which showed that the specific heat diverges, and how the speed of divergences depends on the stimulus ensemble—is largely consistent with a simple, feedforward phenomenological model of retinal processing. Thus, at least for flat models, ‘being very critical’ is a consequence of ‘being strongly correlated’, and not evidence for correlations being fine-tuned or self-organized to a particular value.

Random subsampling gives rise to criticality-inducing correlations

In the above, we showed that a beta-binomial spike-count distribution can be sufficient for signatures of criticality to arise. For this to hold we need the variance of the population spike-count to grow quadratically with population size, i.e. $\text{Var}(K) \propto n^2$. The variance of the population spike-count is equal to the sum of all variances and covariances in the population, $\text{Var}(K) = \sum_{i=1}^n \text{Var}(x_i) + \sum_{i \neq j} \text{Cov}(x_i, x_j)$. A sufficient condition for signatures of criticality to arise in these models is that the average covariances (and hence correlations) between neurons are independent of n , $\frac{1}{n(n-1)} \sum_{i \neq j} \text{Cov}(x_i, x_j) \approx \text{constant}$ [27, 6, 5]. We refer to correlations with this property as ‘criticality inducing’. One possible criticality-inducing correlation structure are so called ‘infinite range’ correlations: correlation between neurons do not drop off to zero for large spatial distances. In the extreme case of distance-independent correlations ([Fig 5a](#)), adding more and more neurons to a population will not change the average pairwise correlation within the population ([Fig 5b](#)). We note that infinite-range correlations are typically not present in the thermodynamic limit in physical systems at equilibrium. In neural systems, infinite-range correlations could be a consequence of densely connected circuitry, or of a shared stimulus drive.

Importantly, criticality-inducing correlations can also result as a consequence of subsampling a large neural population: Even a neural population which does not have infinite-range correlations can appear critical if it is randomly subsampled during analysis. If different populations of size n are obtained as above by (uniformly) subsampling a large recording of size N , then the pairwise correlations in each subpopulation are also a random subsample of the large correlation matrix of the full recording. For any correlation structure on the full recording (including limited-range correlations, [Fig 5c](#)), the expected average correlation in a population of size n is identical to the average correlation in the full recording and hence independent of n ([Fig 5d](#) left, grey line). Despite the pairwise correlations being subsampled in blocks of principal submatrices rather than independently, the variance of the average correlation can drop with the square of the population size n , and is guaranteed to fall at least as $1/n$ (section 4.1 in [S1 Supporting Information](#), and Fig. J in [S1 Supporting Information](#)). Because the average correlation will be independent of n and have negligible variance ([Fig 5d](#) left, shaded area), specific heat will diverge with constant slope ([Fig 5d](#) right). In contrast, if different population sizes are constructed by taking into account the spatial structure of the population (i.e. by

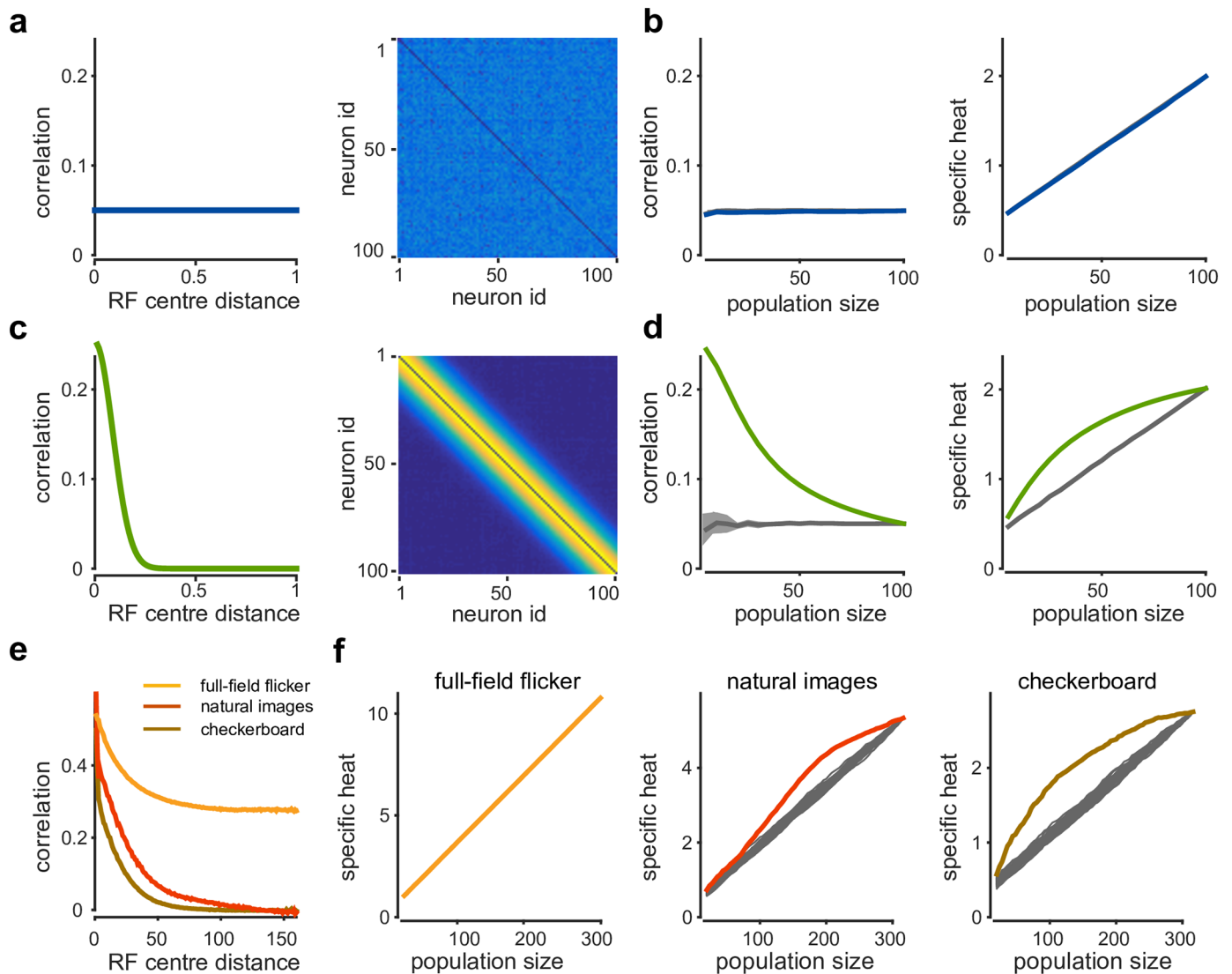


Fig 5. Random subsampling leads to criticality-inducing correlations. **a**) Illustration: A population with 100 neurons and infinite-range correlations, the average correlation between any pair of neurons is close to 0.05. Correlation as function of inter-neuron distance (left) and full correlation matrix (right). **b**) Average correlation in subpopulation of different size n (left) and specific heat at $T = 1$ as function of n (right), when neurons are sampled from 1 to 100 (blue). Random sampling gives identical results (gray). **c**) Population with limited-range correlations, same plots as in panel a. **d**) Left: Average correlation as function of population size for spatially structured sampling (green) and uniform subsampling (gray). Right: Specific heat at $T = 1$ grows linearly for random subsampling, but shows signs of saturation for spatially structured sampling. **e**) Average correlation as function of inter-neuron distance in RGC simulation. For checkerboard and natural images, correlations drop to 0 for large distances. **f**) Specific heat at $T = 1$ for different stimulation conditions, for spatially structured (colour) or random subsampling (gray).

<https://doi.org/10.1371/journal.pcbi.1005718.g005>

iteratively adding neighbouring cells) then the average correlation in each subpopulation will drop with n , and the slope of specific heat growth will decrease with population size.

In our RGC simulation, pairwise correlations did drop off to zero with spatial distance for checkerboard and natural images, but not for full-field flicker (Fig 5e). Pairwise correlations in the full-field flicker condition initially drop off due to distance-dependent shared noise, but eventually saturate at a level far above zero that is determined by the full-field stimulus. Due to these strong infinite-range correlations, both spatially structured sampling and uniform sampling then give rise to linear growth in specific heat (Fig 5f left). For the other stimulus

conditions, however, the choice of subsampling scheme does result in markedly different behavior of the specific heat growth: Both for natural images and checkerboard stimuli, we can see the rate of growth decreases for large n under spatially structured subsampling (Fig 5f centre and right). This effect will be more pronounced for larger simulations, and in additional simulations we found specific heat to saturate once populations are substantially bigger than the spatial range of correlations. This behavior is not unique to the simplified flat models. Specific heat traces computed from K-pairwise models fit to populations obtained with spatially structured sampling also show a marked decrease in specific heat growth rates (section 4.2 in [S1 Supporting Information](#) and Fig. K in [S1 Supporting Information](#)).

In summary, populations will exhibit critical behaviour if correlations have infinite range (over the size of the recording), irrespective of the sampling scheme. In addition, if a population is randomly subsampled (as was done in [7, 8]), then signatures of criticality will arise even if the underlying correlations have limited range.

Discussion

An intriguing hypothesis about the collective activity of large neural populations has been the idea that their statistics resemble those of physical systems at a critical point. In recent years, several studies [12, 5, 6, 11, 7, 8, 18] proposed a new approach to studying criticality in biological data, motivated by notions of criticality in thermodynamics. Signatures of criticality have also been observed in natural images [11] and cortical populations [6], and have been studied using the theory of finite-size scaling and critical exponents [6]. It has been argued that systems close to a critical point might be optimally sensitive to external perturbations [6] and that the large dynamic range of the code (i.e. large variance of log-probabilities) might be beneficial for encoding sensory events which likewise have a large distribution of occurrence probabilities [16].

This hypothesis that neural systems are poised at a thermodynamic critical point could open up further questions on how the system maintains its critical state and on implications for how neural populations encode sensory information and perform computations on it. Alternatively, generic mechanisms could be sufficient to give rise to data which satisfies the definition of criticality put forward in these studies. We had demonstrated in a previous theoretical study [25] that simple models with Gaussian common input can exhibit a diverging specific heat. More recently, it was shown [26, 27, 28] that common input (or other latent variables which lead to shared modulations in firing rates, such as non-stationarity [29]) can give rise to Zipf-like scaling of pattern probabilities, a second signature of criticality. Mathematically, Zipf's Law is equivalent to stating that the plot of entropy vs energy (i.e. log-probability) is a straight line with unit slope [26, 27]. Schwab et al [26] showed that particular latent variable models give rise to Zipf's law. This result was generalized [27, 28] to show that, under fairly general circumstances, high-dimensional latent variable models exhibit a wide distribution of energies (i.e. log-probabilities) and hence a large specific heat. It has also been argued that the use of data sets which are too small might give rise to spuriously big specific heats [49]: while this could be true in principle, additional analyses e.g. in [7] show that their results are robust with respect to data set size, and our results are also valid even in the case of infinite data. Finally, it has also been suggested that whether statistical models exhibit criticality depends on which variables are measured and constrained by the model fit [50, 51].

Previously, criticality in neural systems has also been investigated extensively using a definition of criticality which is based on temporal dynamics with power-law statistics, so-called 'avalanches' [52, 5]. Numerous studies have reported and studied 'avalanche criticality' [8, 21], proposed possible mechanisms (e.g. based on self-organization [53]), and discussed finite-size effects and sub-sampling [54], as well as a need for rigorous statistical analysis [55]. We

emphasize that the ‘avalanche’ definition of criticality is not equivalent to the thermodynamics-inspired definition used in these more recent studies [12, 8]. Our study is only concerned with this more recent approach, and our results thus have no bearing on studies of ‘avalanche-criticality’.

We here related signatures of criticality to the structure of firing rates and correlations in the population: We found that average correlations which are independent of population size are sufficient for inducing criticality, irrespective of their origin. In the thermodynamic analysis of physical systems at equilibrium, long-range correlations typically vanish in the thermodynamic limit. In neural systems, however, ‘criticality-inducing’ correlations can arise as a consequence of various factors: First, in a local patch of retina, retinal ganglion cells have a large degree of receptive field overlap, and natural stimuli also contain strong spatial correlations. This can lead to correlations which do have unlimited range within the experimentally accessible length scales. Thus, fluctuations in the stimulus will lead to common activity modulations amongst neurons within the population. Empirically, correlations between pairs of retinal ganglion cells only fall off slowly with the distance between somata (or receptive field centres) [35]. Second, firing rates e.g. of cortical neurons are modulated by global fluctuations in excitability [45, 56], resulting in neural correlations with infinite range. Third, and importantly, we showed that criticality-inducing correlations can also arise as a consequence of data analysis choices: Uniformly subsampling a large recording with correlations to construct subpopulations yields criticality-inducing correlations, even if the correlations itself do not have unlimited range.

We also showed that there is a direct relationship between ‘how critical’ and ‘how correlated’ a population is: The stronger correlations are, the more prominent the divergence in specific heat is. Mechanisms underlying correlations in spiking activity have been extensively studied in neuroscience [30, 31], and our study makes it possible to relate ‘signatures of criticality’ derived from thermodynamics to these studies, and to interpret the significance of observing these effects: Given the ubiquity of criticality-inducing correlations, signatures of criticality are likely going to be found not just in retinal ganglion cells, but in multiple brain areas and model systems. They are entirely consistent with canonical properties of neural population activity, and require neither finely-tuned parameters in the population, nor sophisticated circuitry or active mechanisms for keeping the system at the critical point. The relationship between firing rates, correlations and criticality (eqs 3 and 4) also yields a prediction about how adaptation in a classical sense should modulate signatures of criticality: The height of the peak is monotonically related to both correlation strength and firing rate. Adaptation typically reduces firing rates and correlations [57, 58]. Taken together, this leads to the prediction that adaptation should *reduce* signatures of criticality— this is precisely the opposite of what has been predicted in [7]. Finally, the dependence of specific heat on correlations might also be an explanation of why Ioffe and Berry [18] found that a feed-forward model fit to their retinal data (which had lower correlations) underestimated the specific heat.

In summary, we conclude that current attempts to interpret findings of thermodynamic criticality in neural population activity have limited potential to lead to new insights into theories of neural computation— in particular, they are not able to discriminate between different hypotheses about either the origin or the functional consequence of the statistics of neural activity. A reliable interpretation of any test for criticality is possible only in reference to a-priori knowledge about the outcome of the test on relevant ground truth models. In order to realise the potential of large-scale recordings of neural activity in the search of a theory of neural computation, we will need data analysis methods which are adapted to the specific properties of biological data, and in particular the fact that neural activity is highly subsampled [59, 60,

54, 61]. One approach to dealing with subsampled data is to use latent-variable models which explicitly model the effect of unobserved inputs and states [62, 63]. In addition, we will also require hypotheses about the normative principles which govern their computations. A possible link between neural activity and theories of criticality might emerge from recent work in machine learning, which is starting to study links between the information-processing capabilities of artificial neural networks and critical phenomena [64].

Materials and methods

Retina simulation

We simulated a population of $N = 316$ retinal ganglion cells as linear threshold neurons whose receptive fields were modelled by difference-of-Gaussian filters with ON-centres [37, 35, 33]. The simulation comprised two subgroups of cells with different receptive field sizes (surrounds $56\mu\text{m}$ and $30\mu\text{m}$ in retinal space, centres $28\mu\text{m}$ and $15\mu\text{m}$, respectively, one third cells with large receptive fields). For both subgroups, the weight of the surround was 0.5 of the centre weight. Locations of receptive field centres (Fig 1 left panel) were based on a reconstruction of 518 soma locations from a patch of mouse retina [65]. As the reconstructed locations in that data set also comprised about 40% amacrine cell somata, we randomly discarded 40% of the cell locations. The resulting patch of retina covered an area of $200 \times 300\mu\text{m}^2$, corresponding to 100×150 pixels in stimulus space. Correlated noise across neurons was modelled using correlated additive Gaussian noise. Correlations dropped off exponentially with soma distance with a decay constant of $\tau = 30\mu\text{m}$ i.e. noise covariance matrix was chosen as

$\Sigma = \sigma_{noise}^2 (aI_n + be^{-D/\tau})$, where D_{ij} is the distance between neurons i and j and $a^2 + b^2 = 1$. We set $\sigma_{noise} = 0.022$ and $a = 0.45$. We modelled neural spiking in discrete time using 20ms bins. In each bin t , the total input $z_i(t)$ to neuron i was given by $z_i(t) = w_i^\top s(t) + \epsilon_i(t)$, where w_i is the receptive field of neuron i , $s(t)$ the vectorised stimulus and $\epsilon_i(t)$ the input noise of neuron i . A neuron in a given bin is active ($x_i = 1$) if $z_i + d > 0.5$ and inactive ($x_i = 0$) otherwise, with offset $d = 0.168$ [66]. Parameters of the simulation (centre and surround sizes, relative strength of centre and surround, magnitude and correlations of noise, spiking threshold) were chosen to roughly match the statistics of neural spiking (firing rates, pairwise correlations, population activity counts) reported in studies of salamander retinal ganglion cells [13, 3, 2].

Stimuli

We used three types of stimuli for this study: natural images, checkerboard patterns and full-field flicker. For natural image stimuli, we used a sequence of 101 images of foliage. Each image was 400×400 pixels, and each image was presented for 20ms with 300 repetitions total. The luminance histograms of the images were transformed to a normal distribution with mean 0.5 and pixel values between 0 and 1.

For the full-field flicker stimulus, luminance levels were drawn from a Gaussian distribution with mean $\mu = 0.5$ and variance $\sigma^2 = 0.06$. Checkerboard stimuli consisted of 80×80 tiles of size 5×5 pixels each. Luminance levels (from within the interval $[0, 1]$) of each tile were chosen to be either 0.15 or 0.77 with probability 0.5. The parameters of both stimulus sets were chosen to match the dynamic range of the simulated retinal ganglion cells. For both types of stimuli, 2000 images were generated and the image sequences were presented with 10 repetitions. To calculate specific heat as function of increasing population size, we randomly selected 10 subsamples of the full simulated population of $N = 316$ cells at population sizes $n \in \{20, 40, 60, 80, 100, 120\}$ by uniformly drawing n neurons out of the full population without replacement.

Statistical model

We modelled retinal ganglion cell activity by using a ‘K-pairwise’ maximum entropy model [3]. In a maximum entropy model [67], the probability of observing the binary spike word $\mathbf{x} \in \{0, 1\}^n$ for parameters $\lambda = \{h, J, V\}$ is given by

$$P(\mathbf{x}|\lambda) = \frac{1}{Z(\lambda)} \exp \left(h^\top \mathbf{x} + \mathbf{x}^\top J \mathbf{x} + \sum_{k=0}^n V_k \delta(K(\mathbf{x}) = k) \right) \quad (5)$$

Here, the parameter vector h (of size $n \times 1$) and the upper-triangular matrix $J \in \mathbb{R}^{n \times n}$ correspond to the bias terms and interaction terms in a pairwise maximum entropy model (also known as an Ising model or spin-glass) [13]. The term $K(\mathbf{x}) = \sum_{i=1}^n x_i$ denotes the population spike-count, i.e. the total number of spikes across the population within a single time bin, and the indicator-term $\delta(K = k)$ is 1 whenever the population spike-count equals k , and is 0 otherwise. The term $\sum_{k=0}^n V_k \delta(K = k)$ was introduced [3] to ensure that the model precisely captures the population spike-count distribution of the data using n additional free parameters. The partition function $Z(\lambda)$ is chosen such that the probabilities of the model sum to 1.

Parameter fitting

To fit the model parameters $\lambda = \{h, J, V\}$ to a data set, we maximised the penalised log-likelihood [68, 69] of the data $\mathcal{D} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(M)}\}$ under the model,

$$L(h, J, V) : = \sum_{m=1}^M \log P(\mathbf{x}^{(m)}|h, J, V) - \frac{1}{\sigma_h} \|h\|_1 - \frac{1}{\sigma_j} \|J\|_1 - \frac{1}{2} V^T \Sigma^{-1} V. \quad (6)$$

Here, the l_1 -penalty controlled the magnitudes of parameters h, J , the term $\|J\|_1$ favoured sparse coupling matrices, and the regularisation term Σ on the V -parameters ensures that the terms controlling the spike-count distribution vary smoothly in k (section 1 in [S1 Supporting Information](#)). This smoothness prior is particularly important for large spike counts, as it makes it possible to interpolate parameters for which the number of observed counts is small.

In maximum entropy models, exact evaluation of the penalised log-likelihood and its gradients requires the calculation of expectations under the model, $E[x_i]$, $E[x_i x_j]$ or equivalently $cov(x_i, x_j)$, and $P(K = k)$ (section 1.1 in [S1 Supporting Information](#)), which in turn requires summations over all 2^n possible states \mathbf{x} and is prohibitive for $n > 20$. Following previous work [15], we used Gibbs sampling to approximate the relevant expectations (section 1.1 in [S1 Supporting Information](#) for derivations and implementation details). We used two modifications over previous applications of Gibbs sampling to fitting maximum entropy models to neural population spike train data, with the goals of speeding up parameter learning and alleviating memory usage:

First, we use Rao-Blackwellisation [40] to speed up convergence of the estimation of covariances of \mathbf{x} : for this, we used pairwise Gibbs sampling (blocked Gibbs with block size 2), where each new sample in the MCMC chain was obtained by updating two entries i and j of \mathbf{x} at a time, rather than just a single entry. This allowed us to get estimates of the conditional probabilities $P(x_i x_j = 1 | x_{\sim \{i,j\}})$, and to use them to speed up the estimation of the second moment $E[x_i x_j]$ from empirical average of these conditional probabilities (section 1.1 in [S1 Supporting Information](#)).

Second, we used a variant of coordinate ascent that calculated all relevant quantities as running averages over the MCMC sample, and thereby avoided having to store the entire $n \times \tilde{M}$ MCMC sample in memory [15], where \tilde{M} is the length of the sample. Because all features of the maximum entropy model are either 0 or 1 ($x_i, x_i x_j$ and the indicator function for the spike count), the gain in log-likelihood obtainable from either updating a single element of h or J [15, 39], or from updating all V simultaneously (but not from updating multiple entries of h and J) can be computed directly from MCMC estimates of $E[x_i]$, $E[x_i x_j]$ and $P(K = k)$ (section 1.2 in S1 Supporting Information). For each iteration, we calculated the gain in log-likelihood for each possible update of h_i, J_{ij} and full V , and picked the update which led to the largest gain [15].

We measured the length of Markov chains in sweeps, where one sweep corresponds to one round of $n(n - 1)/2$ Markov chain updates that encompasses all pairs of entries of \mathbf{x} in random order. We set a learning schedule that started at 800 sweeps for the first parameter update and doubled the number of sweeps in the chain after each set of 1000 parameter updates. We monitored convergence of the algorithm using a normalised mean square error between empirical $E[x_i], cov(x_i, x_j), P(K = k)$ and their estimates from the MCMC sample. For normalisation, we used the average squared values of the target quantity, e.g. $\frac{1}{n} \sum_{i=1}^n E[x_i]^2$ for the firing rates. We stopped the algorithm when a pre-set threshold was reached (0.01%, 0.25%, 0.01% for $E[x_i], cov(x_i, x_j), P(K = k)$, respectively), or when the fitting algorithm took more than $(\frac{n}{100})^2 \times 72\text{h}$ of computation time on a single core (2.294 GHz AMD Opteron(TM) Processor 6276) (Fig. A in S1 Supporting Information). For 10 populations of size $n = 100$ (for natural images), the normalised MSEs after model-fitting were 0.43%, 2.80%, 0.42%). An implementation of the fitting algorithms in MATLAB is available at <https://github.com/mackelab/CorBinian>.

Specific heat calculation

To investigate thermodynamic properties of neural population codes, Tkačik et al [7] introduced a temperature parameter T for eq 5:

$$P_T(\mathbf{x}|\lambda) = \frac{1}{Z_T} \exp \left(\frac{1}{T} \left(h^\top \mathbf{x} + \mathbf{x}^\top J \mathbf{x} + \sum_{k=0}^n V_k \delta(K(\mathbf{x}) = k) \right) \right) \tag{7}$$

Model fits are obtained at $T = 1$, and the temperature parameter T is scaled to study the system (i.e. characterised by $P_T(\mathbf{x}|h, J, V)$ for $T = 1$). Varying T , in effect, modulates probabilities by exponentiating them with $1/T$,

$$P_T(\mathbf{x}) \propto (P_{T=1}(\mathbf{x}))^{1/T}, \tag{8}$$

and that the family of probability distributions obtained by varying T can be constructed for any distribution, not just maximum entropy models. For large temperatures P_T approaches a uniform distribution ($P_T(\mathbf{x}) \approx 2^{-n}$ for each \mathbf{x}), whereas for small temperatures it converges to a singleton, $P_T(\mathbf{x}^*) \approx 1$ with $\mathbf{x}^* = \text{argmax}_{\mathbf{x}}(P_{T=1}(\mathbf{x}))$.

The specific heat, as given in eq 2, can be obtained from the variance of the log-probabilities of the model. As the variance in practice cannot be computed for large n , we obtained estimates of $c(T)$ using a pairwise Gibbs sampler. The specific heat does not depend on Z_T , as changing Z_T results in a constant, additive shift in log-probabilities which does not affect the variance. We tracked the variance of log-probabilities over an MCMC chain $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\tilde{M})}$ of

length \tilde{M} sampled at temperature T , using

$$c(T) \approx \frac{1}{n} (\hat{E}[\log P_T(\mathbf{x}^{(m)}|\lambda)^2] - \hat{E}[\log P_T(\mathbf{x}^{(m)}|\lambda)]^2) \tag{9}$$

where \hat{E} denotes the average over spike words $\mathbf{x}^{(m)}$ sampled from the the MCMC chain. For each population, we evaluated $c(T)$ for 31 temperatures between $T = 0.8$ and $T = 2$, and found the Gibbs sampler to provide reliable estimates over this temperature range—we in particular chose the minimal temperature $T = 0.8$ larger than previous previously in [7] to minimize possible effects from the sampler getting stuck (see e.g. [46]). We used a burn-in of 2.0e4 sweeps, and ran the sampler for $(\frac{n}{100})^2 \times 4$ h of CPU time, resulting in between 9.97e5 and 1.72e6 sweeps for $n = 100$ (i.e. between 4.94e9 and 8.52e9 sampled individual spike words).

Simplified population models

For the theoretical analysis of the sampling process, we adopted a class of population models (here referred to as ‘flat’ models) in which all neurons are drawn from an infinite pool of neurons which all have identical mean firing rates, pairwise correlations and higher-order correlations [44, 25, 70, 3, 71]. Such a model is fully specified by the population spike-count distribution $P(K = k)$, and all spike words with the same spike count are equally probable. As a result, the probabilities of individual patterns \mathbf{x} can be read off from the spike-count distribution by

$$P(\mathbf{x}) = \binom{n}{k}^{-1} P(K = k) \tag{10}$$

whenever $\sum_{i=1}^n x_i = k$. In a maximum entropy formalism, this model can be obtained by setting $h_i = 0$ and $J_{ij} = 0$ for all $i, j \in \{1, \dots, n\}$ and only optimising entries of V . Without loss of generality, we fixed $V_0 = 0$ [43], resulting in n degrees of freedom for the model.

In flat models, it is possible to explicitly construct a limit $n \rightarrow \infty$ which will help us understand population analyses performed on experimental data: We assume that there is a spike-count density $f(r)$, $r \in [0, 1]$, which describes the population spike-count distribution of an infinitely large population. $f(r)$ denotes the probability density of a fraction of r neurons spiking simultaneously. Finite-size populations of n cells are then obtained as random subsamples out of this infinitely large system. Based on previous findings by [25], we show in section 2.3 in [S1 Supporting Information](#) that, in this construction, flat models always exhibit a linear divergence of specific heat, unless the limit $f(r)$ is given by either a single delta peak or a mixture of two symmetric delta peaks. These two models corresponds to systems that (for large n) either behave like a fully independent population (whose spike-count distribution converges to a single delta peak), or a population described by a pure pairwise maximum entropy model (which converges to two delta peaks). In particular, any flat model with higher-order correlations [17, 70, 71], or a non-degenerate $f(r)$, will exhibit ‘signatures of criticality’. Furthermore, we show that, for continuous $f(r)$, $c(T)$ does not diverge for any $T \neq 1$. In combination, these results show that the peak of the specific heat is mathematically bound to converge to $T = 1$ for $n \rightarrow \infty$ in this model class.

We further simplified the flat model by re-parametrising $P(K = k)$ by a beta-binomial distribution, thereby reducing the number of parameters from n to two, and—importantly—obtaining parameters which do not explicitly depend on n . In this model,

$$P(K = k) = \binom{n}{k} \frac{\text{Beta}(\alpha + k, \beta + n - k)}{\text{Beta}(\alpha, \beta)} = \binom{n}{k} \int f(r) r^k (1 - r)^{n-k} dr \tag{11}$$

and

$$f(r) = \frac{1}{\text{Beta}(\alpha, \beta)} r^{\alpha-1} (1-r)^{\beta-1}. \quad (12)$$

For simulated data, we found values for α, β extracted from the beta-binomial fits to populations of different sizes n to be stable over a large range of n (Fig 3b). We used the beta-binomial parameters obtained from the largest investigated n to estimate the divergence rate \tilde{c} for $n \rightarrow \infty$.

Supporting information

S1 Supporting Information. Supporting derivations and analyses. We provide more detailed descriptions of the maximum entropy fitting procedures used in this study. We derive limiting behavior of specific heat capacity for flat models, and analyze effects of uniform subsampling on sample means and variances. Furthermore, we provide control analyses for central findings of the study.

(PDF)

Acknowledgments

We thank F. Franzen for help with figures and cluster computing, and S. Buchholz, D. Greenberg, S. Turaga and A. Renart for discussions and comments on the manuscript.

Author Contributions

Conceptualization: Marcel Nonnenmacher, Jakob H. Macke.

Data curation: Marcel Nonnenmacher, Christian Behrens, Philipp Berens.

Formal analysis: Marcel Nonnenmacher, Jakob H. Macke.

Funding acquisition: Philipp Berens, Matthias Bethge, Jakob H. Macke.

Investigation: Marcel Nonnenmacher, Jakob H. Macke.

Methodology: Marcel Nonnenmacher, Christian Behrens, Jakob H. Macke.

Project administration: Jakob H. Macke.

Resources: Matthias Bethge, Jakob H. Macke.

Software: Marcel Nonnenmacher, Christian Behrens, Jakob H. Macke.

Supervision: Philipp Berens, Matthias Bethge, Jakob H. Macke.

Validation: Marcel Nonnenmacher, Philipp Berens, Jakob H. Macke.

Visualization: Marcel Nonnenmacher, Jakob H. Macke.

Writing – original draft: Marcel Nonnenmacher, Jakob H. Macke.

Writing – review & editing: Marcel Nonnenmacher, Christian Behrens, Philipp Berens, Matthias Bethge, Jakob H. Macke.

References

1. Kerr JND, Denk W. Imaging in vivo: watching the brain in action. *Nature Reviews Neurosci.* 2008; 9(3):195–205. <https://doi.org/10.1038/nrn2338>

2. Marre O, Amodei D, Deshmukh N, Sadeghi K, Soo F, Holy TE, et al. Mapping a complete neural population in the retina. *The Journal of Neuroscience*. 2012; 32(43):14859–14873. <https://doi.org/10.1523/JNEUROSCI.0723-12.2012> PMID: 23100409
3. Tkačik G, Marre O, Amodei D, Schneidman E, Bialek W, Berry MJ 2nd. Searching for collective behavior in a large network of sensory neurons. *PLoS Comput Biol*. 2014; 10(1):e1003408. <https://doi.org/10.1371/journal.pcbi.1003408> PMID: 24391485
4. Gao P, Ganguli S. On simplicity and complexity in the brave new world of large-scale neuroscience. *Current opinion in neurobiology*. 2015; 32:148–155. <https://doi.org/10.1016/j.conb.2015.04.003> PMID: 25932978
5. Beggs JM, Timme N. Being critical of criticality in the brain. *Frontiers in physiology*. 2012; 3. <https://doi.org/10.3389/fphys.2012.00163> PMID: 22701101
6. Yu S, Yang H, Shriki O, Plenz D. Universal organization of resting brain activity at the thermodynamic critical point. *Front Syst Neurosci*. 2013; 7:42. <https://doi.org/10.3389/fnsys.2013.00042> PMID: 23986660
7. Tkačik G, Mora T, Marre O, Amodei D, Palmer SE, Berry MJ, et al. Thermodynamics and signatures of criticality in a network of neurons. *Proceedings of the National Academy of Sciences*. 2015; 112(37):11508–11513. <https://doi.org/10.1073/pnas.1514188112>
8. Mora T, Deny S, Marre O. Dynamical criticality in the collective activity of a population of retinal neurons. *Physical review letters*. 2015; 114(7):078105. <https://doi.org/10.1103/PhysRevLett.114.078105> PMID: 25763977
9. Mora T, Walczak AM, Bialek W, Callan CG. Maximum entropy models for antibody diversity. *Proceedings of the National Academy of Sciences*. 2010; 107(12):5405–5410. <https://doi.org/10.1073/pnas.1001705107>
10. Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M, et al. Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences*. 2012; 109(13):4786–4791. <https://doi.org/10.1073/pnas.1118633109>
11. Stephens GJ, Mora T, Tkačik G, Bialek W. Statistical thermodynamics of natural images. *Phys Rev Lett*. 2013 Jan; 110(1):018701. <https://doi.org/10.1103/PhysRevLett.110.018701> PMID: 23383852
12. Mora T, Bialek W. Are biological systems poised at criticality? *Journal of Statistical Physics*. 2011; 144(2):268–302.
13. Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*. 2006; 440(7087):1007–12. <https://doi.org/10.1038/nature04701> PMID: 16625187
14. Shlens J, Field GD, Gauthier JL, Grivich MI, Petrusca D, Sher A, et al. The structure of multi-neuron firing patterns in primate retina. *J Neurosci*. 2006; 26(32):8254–66. <https://doi.org/10.1523/JNEUROSCI.1282-06.2006> PMID: 16899720
15. Broderick T, Dudik M, Tkacik G, Schapire RE, Bialek W. Faster solutions of the inverse pairwise Ising problem. *arXiv*. 2007;0712.2437v2.
16. Tkacik G, Schneidman E, Berry MJ II, Bialek W. Spin glass models for a network of real neurons. *arXiv: q-bio/0611072v2*. 2009.
17. Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*. 2010; 466(7306):617–621. <https://doi.org/10.1038/nature09178> PMID: 20601940
18. Ioffe ML, Berry II J Michael. The StructuredLow Temperature Phase of the Retinal Population Code. *arXiv preprint arXiv:160805751*. 2016.
19. Hosoya T, Baccus SA, Meister M. Dynamic predictive coding by the retina. *Nature*. 2005; 436(7047):71. <https://doi.org/10.1038/nature03689> PMID: 16001064
20. Prentice J, Simmons K, Tkacik G, Homann J, Yee H, Palmer S, et al. Transformation of stimulus correlations by the retina. In: *APS Meeting Abstracts*. vol. 1; 2014. p. 15002.
21. Shew WL, Clawson WP, Pobst J, Karimipناه Y, Wright NC, Wessel R. Adaptation to sensory input tunes visual cortex to criticality. *Nature Physics*. 2015; 11(8):659–663. <https://doi.org/10.1038/nphys3370>
22. Bak P, Tang C, Wiesenfeld K. Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*. 1987; 59(4):381. <https://doi.org/10.1103/PhysRevLett.59.381> PMID: 10035754
23. Levina A, Herrmann JM, Geisel T. Phase transitions towards criticality in a neural system with adaptive interactions. *Physical Review Letters*. 2009; 102(11):118110. <https://doi.org/10.1103/PhysRevLett.102.118110> PMID: 19392248
24. Meisel C, Gross T. Adaptive self-organization in a realistic neural network model. *Physical Review E*. 2009; 80(6):061917. <https://doi.org/10.1103/PhysRevE.80.061917>

25. Macke JH, Opper M, Bethge M. Common input explains higher-order correlations and entropy in a simple model of neural population activity. *Physical Review Letters*. 2011; 106(20):208102. <https://doi.org/10.1103/PhysRevLett.106.208102> PMID: 21668265
26. Schwab DJ, Nemenman I, Mehta P. Zipf's law and criticality in multivariate data without fine-tuning. *Physical review letters*. 2014; 113(6):068102. <https://doi.org/10.1103/PhysRevLett.113.068102> PMID: 25148352
27. Aitchison L, Corradi N, Latham PE. Zipf's law arises naturally in structured, high-dimensional data. *arXiv preprint*. 2014;1407.7135v4.
28. Aitchison L, Corradi N, Latham PE. Zipf's Law Arises Naturally When There Are Underlying, Unobserved Variables. *PLoS Comput Biol*. 2016 Dec; 12(12):e1005110. <https://doi.org/10.1371/journal.pcbi.1005110> PMID: 27997544
29. Tyrcha J, Roudi Y, Marsili M, Hertz J. The effect of nonstationarity on models inferred from neural data. *Journal of Statistical Mechanics: Theory and Experiment*. 2013; 2013(03):P03005. <https://doi.org/10.1088/1742-5468/2013/03/P03005>
30. Trong PK, Rieke F. Origin of correlated activity between parasol retinal ganglion cells. *Nature Neuroscience*. 2008; 11(11):1343–1351. <https://doi.org/10.1038/nn.2199> PMID: 18820692
31. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K. The mechanics of state-dependent neural correlations. *Nat Neurosci*. 2016 Mar; 19(3):383–93. <https://doi.org/10.1038/nn.4242> PMID: 26906505
32. Chichilnisky E. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*. 2001; 12(2):199–213. <https://doi.org/10.1080/1080713663221>
33. Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, et al. Do we know what the early visual system does? *J Neurosci*. 2005; 25(46):10577–10597. <https://doi.org/10.1523/JNEUROSCI.3726-05.2005> PMID: 16291931
34. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, et al. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*. 2008; 454(7207):995–9. <https://doi.org/10.1038/nature07140> PMID: 18650810
35. Pitkow X, Meister M. Decorrelation and efficient coding by retinal ganglion cells. *Nature neuroscience*. 2012; 15(4):628–635. <https://doi.org/10.1038/nn.3064> PMID: 22406548
36. Kuffler SW. Discharge patterns and functional organization of mammalian retina. *Journal of neurophysiology*. 1953; 16(1):37–68. PMID: 13035466
37. Rodieck RW. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision research*. 1965; 5(12):583–601. [https://doi.org/10.1016/0042-6989\(65\)90033-7](https://doi.org/10.1016/0042-6989(65)90033-7) PMID: 5862581
38. Ferrenberg AM, Swendsen RH. New Monte Carlo technique for studying phase transitions. *Physical review letters*. 1988; 61(23):2635. <https://doi.org/10.1103/PhysRevLett.61.2635> PMID: 10039183
39. Schwartz G, Macke J, Amodei D, Tang H, Berry MJ 2nd. Low error discrimination using a correlated population code. *J Neurophysiol*. 2012; 108(4):1069–88. <https://doi.org/10.1152/jn.00564.2011> PMID: 22539825
40. Radhakrishna Rao C. Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*. 1945; 37(3):81–91.
41. Kirkpatrick S, Gelatt CD, Vecchi MP, et al. Optimization by simulated annealing. *science*. 1983; 220(4598):671–680. <https://doi.org/10.1126/science.220.4598.671> PMID: 17813860
42. Tkacik G, Schneidman E, Berry II MJ, Bialek W. Ising models for networks of real neurons. *arXiv preprint*. 2006;0611072v1.
43. Tkačik G, Marre O, Mora T, Amodei D, Berry MJ II, Bialek W. The simplest maximum entropy model for collective behavior in a neural network. *Journal of Statistical Mechanics: Theory and Experiment*. 2013; 2013(03):P03011. <https://doi.org/10.1088/1742-5468/2013/03/P03011>
44. Amari Si, Nakahara H, Wu S, Sakai Y. Synchronous firing and higher-order interactions in neuron pool. *Neural Computation*. 2003; 15(1):127–142. <https://doi.org/10.1162/089976603321043720> PMID: 12590822
45. Okun M, Yger P, Marguet SL, Gerard-Mercier F, Benucci A, Katzner S, et al. Population rate dynamics and multineuron firing patterns in sensory cortex. *J Neurosci*. 2012; 32(48):17108–19. <https://doi.org/10.1523/JNEUROSCI.1831-12.2012> PMID: 23197704
46. Rostami V, Mana PP, Helias M. Pairwise maximum-entropy models and their Glauber dynamics: bimodality, bistability, non-ergodicity problems, and their elimination via inhibition. *arXiv preprint arXiv:160504740*. 2016.
47. Sherrington D, Kirkpatrick S. Solvable model of a spin-glass. *Physical review letters*. 1975; 35(26):1792. <https://doi.org/10.1103/PhysRevLett.35.1792>
48. Mezard M, Parisi G, Virasoro M. *Spin Glass Theory and Beyond* (Singapore: Word Scientific); 1987.

49. Saremi S, Sejnowski TJ. On Criticality in High-Dimensional Data. *Neural Comput.* 2014 Jul; 26(7):1329–1339. https://doi.org/10.1162/NECO_a_00607 PMID: 24708368
50. Mastromatteo I, Marsili M. On the criticality of inferred models. *Journal of Statistical Mechanics: Theory and Experiment.* 2011; 2011(10):P10012. <https://doi.org/10.1088/1742-5468/2011/10/P10012>
51. Marsili M, Mastromatteo I, Roudi Y. On sampling and modeling complex systems. *Journal of Statistical Mechanics: Theory and Experiment.* 2013; 2013(09):P09003. <https://doi.org/10.1088/1742-5468/2013/09/P09003>
52. Beggs JM, Plenz D. Neuronal avalanches in neocortical circuits. *The Journal of neuroscience.* 2003; 23(35):11167–11177. PMID: 14657176
53. Levina A, Herrmann JM, Geisel T. Dynamical synapses causing self-organized criticality in neural networks. *Nature physics.* 2007; 3(12):857–860. <https://doi.org/10.1038/nphys758>
54. Levina A, Priesemann V. Subsampling scaling. *Nature Communications.* 2017; 8:15140. <https://doi.org/10.1038/ncomms15140> PMID: 28469176
55. Touboul J, Destexhe A. Power-law statistics and universal scaling in the absence of criticality. *Physical Review E.* 2017; 95(1):012413. <https://doi.org/10.1103/PhysRevE.95.012413> PMID: 28208383
56. Schölvinck ML, Saleem AB, Benucci A, Harris KD, Carandini M. Cortical state determines global variability and correlations in visual cortex. *J Neurosci.* 2015 Jan; 35(1):170–8. <https://doi.org/10.1523/JNEUROSCI.4994-13.2015> PMID: 25568112
57. Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M, et al. Adaptation of retinal processing to image contrast and spatial scale. *Nature.* 1997; 386(6620):69–73. <https://doi.org/10.1038/386069a0> PMID: 9052781
58. Nirenberg S, Bomash I, Pillow JW, Victor JD. Heterogeneous response dynamics in retinal ganglion cells: the interplay of predictive coding and adaptation. *Journal of neurophysiology.* 2010; 103(6):3184–3194. <https://doi.org/10.1152/jn.00878.2009> PMID: 20357061
59. Turaga S, Buesing L, Packer AM, Dalgleish H, Pettit N, Hausser M, et al. Inferring neural population dynamics from multiple partial recordings of the same neural circuit. In: *Advances in Neural Information Processing Systems*; 2013. p. 539–547.
60. Soudry D, Keshri S, Stinson P, Oh Mh, Iyengar G, Paninski L. Efficient “Shotgun” Inference of Neural Connectivity from Highly Sub-sampled Activity Data. *PLoS Comput Biol.* 2015; 11(10):e1004464. <https://doi.org/10.1371/journal.pcbi.1004464> PMID: 26465147
61. Battistin C, Dunn B, Roudi Y. Learning with unknowns: analyzing biological data in the presence of hidden variables. *Current Opinion in Systems Biology.* 2017. <https://doi.org/10.1016/j.coisb.2016.12.010>
62. Macke JH, Buesing L, Cunningham JP, Yu BM, Shenoy KV, Sahani M. Empirical models of spiking in neural populations. In: *Advances in Neural Information Processing Systems*. vol. 24; 2012.
63. Archer EW, Koster U, Pillow JW, Macke JH. Low-dimensional models of neural population activity in sensory cortical circuits. In: *Advances in Neural Information Processing Systems*; 2014. p. 343–351.
64. Schoenholz SS, Gilmer J, Ganguli S, Sohl-Dickstein J. Deep Information Propagation. *arXiv preprint arXiv:161101232.* 2016.
65. Baden T, Berens P, Franke K, Rosón MR, Bethge M, Euler T. The functional diversity of retinal ganglion cells in the mouse. *Nature.* 2016; 529(7586):345–350. <https://doi.org/10.1038/nature16468> PMID: 26735013
66. Lyamzin DR, Macke JH, Lesica NA. Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations. *Frontiers in computational neuroscience.* 2010; 4. <https://doi.org/10.3389/fncom.2010.00144> PMID: 21152346
67. Jaynes ET. Information theory and statistical mechanics. *Physical review.* 1957; 106(4):620. <https://doi.org/10.1103/PhysRev.106.620>
68. Dudík M, Schapire RE. Maximum entropy distribution estimation with generalized regularization. In: *Learning Theory*. Springer; 2006. p. 123–138.
69. Altun Y, Smola A. Unifying divergence minimization and statistical inference via convex duality. In: *Learning theory*. Springer; 2006. p. 139–153.
70. Yu S, Yang H, Nakahara H, Santos GS, Nikolic D, Plenz D. Higher-order interactions characterized in cortical activity. *J Neurosci.* 2011; 31(48):17514–17526. <https://doi.org/10.1523/JNEUROSCI.3127-11.2011> PMID: 22131413
71. Barreiro AK, Gjorgjieva J, Rieke F, Shea-Brown E. When do microcircuits produce beyond-pairwise correlations? *Front Comput Neurosci.* 2014; 8:10. <https://doi.org/10.3389/fncom.2014.00010> PMID: 24567715