

Supplementary details and explanations for paper “A model of neural population activity in the presence of common input explains higher-order correlations and entropy”

Jakob H Macke,^{1,*} Manfred Opper,² and Matthias Bethge³

¹*Gatsby Unit, University College London, UK, and University of Tübingen, Germany*

²*Artificial Intelligence Group, Technical University Berlin, Germany*

³*Institute for Theoretical Physics, Werner Reichardt Centre for Integrative Neuroscience,
Bernstein Center for Computational Neuroscience,*

MPI for Biological Cybernetics and University of Tübingen, Germany

(Dated: April 19, 2011)

This manuscript contains additional material for the manuscript “A model of neural population activity in the presence of common input explains higher-order correlations and entropy”. It contains very detailed step-by-step derivations and proofs for all the formulas used in the paper. We aimed to present derivations in a self-contained manner, rather than in a brief and elegant one. We also included some additional explanations, additional figures and enlarged versions of some figures, as well as additional references.

I. PROPERTIES OF THE DICHOTOMIZED GAUSSIAN

A. Introduction

We use n -dimensional binary random variable $X \in \{0, 1\}^n$ with mean μ and covariance matrix Σ as a simplified model of neural population statistics. For details on how to fit our common input model, which is equivalent to Dichotomized Gaussian [8, 18] models, and applications to modelling binarized natural images and neural spike-trains, see [3, 12, 13]. In particular, the technical report [13] contains some derivations which are overlapping with the ones in this document.

In the letter, we concentrated on the effects of instantaneous correlations across neurons in binarized spike trains, as these have been most thoroughly characterized in experimental studies [7, 22–25]. Other studies [5, 11] have used doubly-stochastic processes with Gaussian process rate functions to generate spike trains with controlled correlation structure in continuous time. [6, 9, 26] have studied the relationship between input correlations, output correlations and firing rates in threshold neurons receiving temporally and spatially correlated inputs, and have also investigated temporal properties of these models. However, none of these studies focused on the effects of pairwise correlations on population statistics such as entropy, sparsity or heat-capacity. More recently, [2] have investigated the question of when small neural populations receiving common-input are modelled well by pairwise maximum entropy methods. [14] used a second order Maximum Entropy model with additional temporal dependencies (under a Markov assumption) model spatial as well as temporal correlations in neural population recordings. [19–21] studied maximum entropy models of second order and compared them with models containing higher-order interactions by using different expansion methods. In particular, they also derived and validated analytical approximations for inferring the parameters of maximum entropy methods, and characterized the scaling of entropy in the ‘perturbative’ regime.

B. Spike count distribution

In the letter, we concentrate on homogeneous populations with n neurons, each having firing rate μ , constant pairwise covariance σ and correlation coefficient $\rho = \sigma/(\mu(1 - \mu))$ [1, 4, 15]. Binary samples from the model are obtained by dichotomizing a latent Gaussian Z with mean γ , unit variances and pairwise covariance (and correlation coefficient) $\lambda > 0$. This model was introduced and first studied by [1].

*Electronic address: jakob@gatsby.ucl.ac.uk

We set

$$Z \sim \mathcal{N}(\gamma, \Lambda), \text{ where } \Lambda = \mathbb{I}(1 - \lambda) + \lambda e_n e_n^\top \quad (1)$$

$$Z = \gamma + \sqrt{1 - \lambda}T + \sqrt{\lambda}e_n S \quad (2)$$

$$\text{where } T \sim \mathcal{N}(0, \mathbb{I}_n) \quad (3)$$

$$S \sim \mathcal{N}(0, 1), \quad (4)$$

where e_n is a column vector consisting of all ones.

The binary patterns X are generated by thresholding Z at 0, i.e. by setting $X_i = 1$ whenever $Z_i > 0$, and $X_i = 0$ otherwise. Because of symmetry, $P(X = x) = P(X = y)$ whenever $\|x\| = \|y\|$, where $\|x\| = \sum_i x_i$. Such a homogeneous population model is fully characterized by the number of neurons that spike synchronously, i.e. the population spike count distribution: We set $K = \|X\|$, the population spike count, and $R = K/n$, the proportion of neurons spiking at any time. The probability of seeing a given pattern with k spikes is $Q_{DG}(k) = P_{DG}(K = k) / \binom{n}{k}$.

The probability of observing a pattern x is given by

$$P_{DG}(X = x) = \int_z P(X = x|z)p(z)dz \quad (5)$$

$$= \int_{-\infty}^{\infty} \phi_\lambda(s) \prod_{i=1}^n P(X_i = x_i|s)ds \quad (6)$$

$$= \int_{-\infty}^{\infty} \phi_\lambda(s) \prod_{i=1}^n (1 - L(s))^{1-x_i} L(s)^{x_i} ds, \quad (7)$$

$$\text{where } L(s) = P(Z_i > 0|s) \quad (8)$$

$$= \Phi\left(\frac{s + \gamma}{\sqrt{1 - \lambda}}\right). \quad (9)$$

Here, $\phi_\lambda(s)$ is the probability density function of a one-dimensional Gaussian with mean 0 and variance λ . Similarly, the probability of observing a spike count $K = k$ is

$$P_{DG}(K = k) = \binom{n}{k} P(X = x, \text{ where } \|x\| = k) \quad (10)$$

$$= \binom{n}{k} \int_{-\infty}^{\infty} \phi_\lambda(s) (1 - L(s))^{n-k} L(s)^k ds. \quad (11)$$

C. Strain of the DG

[17] defined the strain ψ of a distribution over three binary random variables to be the third-order interaction term in a maximum entropy model of degree three with $\sigma_i = -1, 1$ spins, i.e. $P(\sigma) = \frac{1}{Z} \exp\left(\sum_i h_i \sigma_i + \sum_{i < j} J_{ij} \sigma_i \sigma_j + \gamma \sigma_1 \sigma_2 \sigma_3\right)$. They showed that γ could also be calculated via the formulas

$$\psi = \frac{1}{8} \sum \sigma_1 \sigma_2 \sigma_3 P(\sigma) \quad (12)$$

$$\psi = \frac{1}{8} \log \frac{P_{100} P_{010} P_{001} P_{111}}{P_{000} P_{011} P_{101} P_{110}} \quad (13)$$

The strain ψ measures the correction one has to apply to the probabilities of an second order maximum entropy model to also account for triplet-interactions.

We calculated the strain of a homogeneous DG with three neurons, using formula (13), with log base 2. Because of symmetries, it is sufficient to evaluate P_{000} , P_{001} and P_{011} . Each of these three probabilities can be calculated using a one-dimensional Gaussian integral as in equation (11). Thus, while there is no closed form expression for the DG in general, the strain can be calculated very precisely using numerical integration. For all possible values

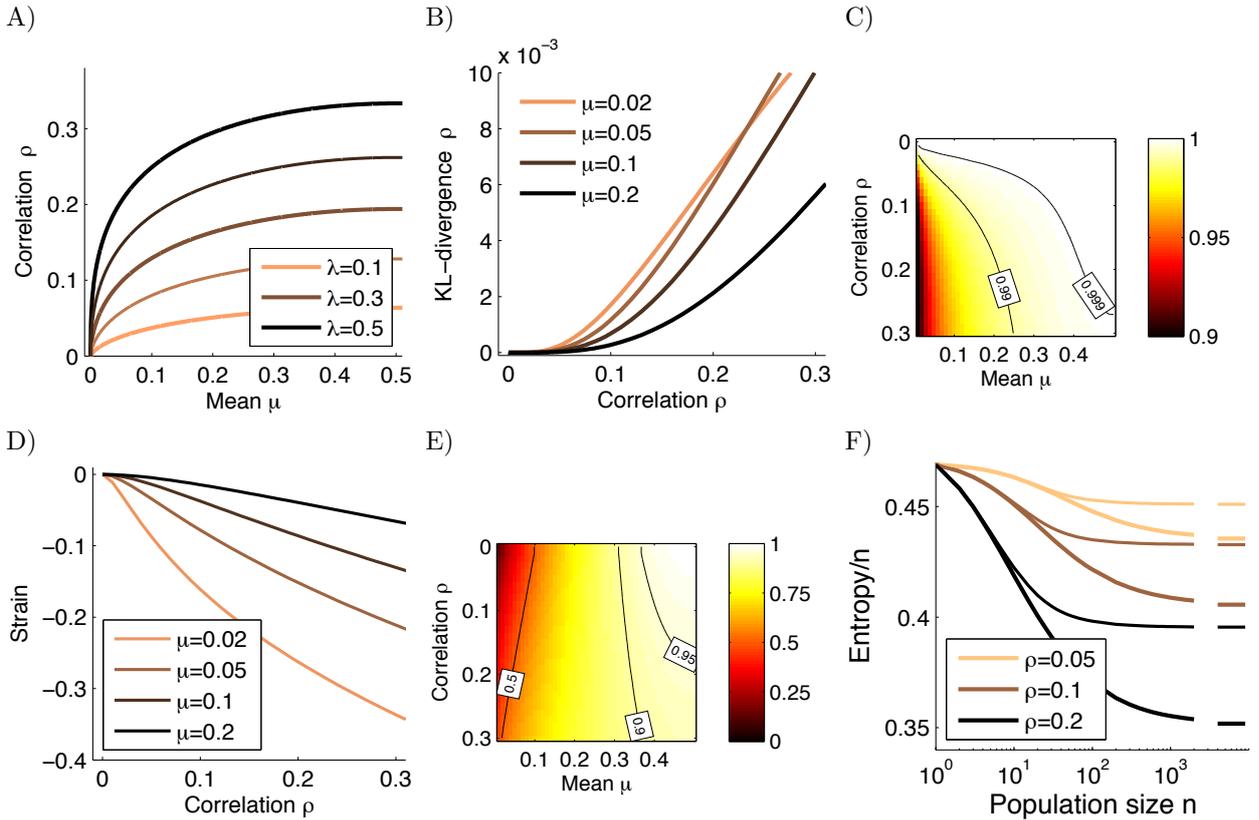


FIG. 1: Correlations in the DG. This figure is identical to figure 1 of the letter, but enlarged for better visibility. **a)** The correlation coefficient ρ increases with firing probability μ for constant input correlation λ . **b)** The KL-divergence Δ_h between the DG and its second order approximation is modulated by the mean firing rate μ and the correlation ρ in a population of size $n = 5$. **c)** In small populations ($n = 5$), the multi-information explained (I_2) by a DG is close to 1. **d)** The strain of the homogeneous DG is negative and correlation-dependent. **e)** For large populations, I_2 between the models close to 0 for small correlations, i.e. the models are very dissimilar. **f)** Scaling of the entropy rate (i.e. entropy per neuron) of the Ising and DG model for mean $\mu = 0.1$, and comparison with asymptotic values (labeled ∞). The entropy rate drops initially before settling to the asymptotic value. For weak correlations, differences between models only become apparent for large n .

of μ ($\mu \leq 0.5$) and ρ , the strain was negative. Similarly, the probability of three simultaneous spikes P_{111} was consistently *lower* in the DG than in the MaxEnt model. The homogeneous MaxEnt model actually over-estimates the occurrence-probabilities of three simultaneous spikes in this common input model.

To test whether this also holds true for more general correlation structures, we simulated random networks of three neurons by sampling the parameter γ from a three-dimensional Gaussian (and flipping the output to ensure $\gamma < 0$, i.e. $\mu < 0.5$), and sampling random covariance matrices Λ by defining Λ to be the product of a randomly sampled 3×3 matrix with its transpose. We varied the means and variances, with the goal of sampling the space of all admissible means and covariances for a three dimensional DG. In the vast majority of cases in which all three pairwise correlations had the same sign, the strain was negative. Thus, a negative strain is not necessarily inconsistent with common input being the source of higher-order redundancies. See [2] for more detailed investigations of the effect of different common input statistics on the strain.

II. ASYMPTOTIC PROPERTIES OF THE DG

We present two derivations of the asymptotic spike count distribution $f(r)$ of the DG. The first uses a saddle-point approximation and is similar to a derivation in [1]. The second is based on the observation that the percentage of neurons that spike simultaneously for a given input must be the same as the probability of a single neuron spiking for that input. Finally, we calculate the entropy and asymptotic entropy rate of the DG-model from its spike-count distributions, as well as the specific heat of the DG at unit temperature.

A. Derivation of spike-count distribution via saddle-point approximation

For large population sizes n , the integral in (11) can be approximated using the saddle-point approximation: We set $U(s) = r \log L(s) + (1-r) \log L(-s)$ to get

$$Q_{DG}(k) = \int_{-\infty}^{\infty} \phi_{\lambda}(s) L(s)^{rn} (1-L(s))^{n-rn} ds \quad (14)$$

$$= \int_{-\infty}^{\infty} \phi_{\lambda}(s) e^{nU(s)} ds \quad (15)$$

$$\text{for large } n \approx \sqrt{\frac{2\pi}{-nU''(s_o)}} \phi_{\lambda}(s_o) e^{nU(s_o)} \quad (16)$$

$U(s)$ has a unique maximum u_o , as it is just a binomial likelihood function in L , and L is monotonic in s . Therefore, we get that

$$Q_{DG}(k) = \frac{\phi(s_o/\sqrt{\lambda})}{\phi\left(\frac{s_o+\gamma}{\sqrt{1-\lambda}}\right)} (r^r(1-r)^{1-r})^n \sqrt{\frac{2\pi(1-\lambda)r(1-r)}{\lambda n}} \quad (17)$$

$$(18)$$

and the discrete 'large n ' population spike count distribution

$$P_{DG}(K = k) = Q_{DG}(k) \binom{n}{k} \quad (19)$$

$$= \frac{\phi(s_o/\sqrt{\lambda})}{\phi(\Phi^{-1}(r))} \sqrt{\frac{2\pi(1-\lambda)r(1-r)}{\lambda}} \left(\frac{(r^r(1-r)^{1-r})^n}{\sqrt{n}} \binom{n}{nr} \right) \quad (20)$$

$$\approx \frac{\phi(s_o/\sqrt{\lambda})}{\phi(\Phi^{-1}(r))} \sqrt{\frac{2\pi(1-\lambda)r(1-r)}{\lambda}} \left(\frac{1}{n\sqrt{2\pi r(1-r)}} \right). \quad (21)$$

Renormalizing, this yields the asymptotic, continuous valued spike count distribution:

$$f(r) = \frac{\phi(s_o/\sqrt{\lambda})}{\phi(\Phi^{-1}(r))} \sqrt{\frac{1-\lambda}{\lambda}} \quad (22)$$

$$= \exp\left(-\frac{1}{2} \frac{(\nu - \frac{\gamma\sqrt{1-\lambda}}{1-2\lambda})^2}{\lambda/(1-2\lambda)}\right) \exp\left(\frac{\gamma^2}{2-4\lambda}\right) \sqrt{\frac{1-\lambda}{\lambda}} \quad (23)$$

where $\nu = \Phi^{-1}(r)$.

Equivalently, we can write

$$f_{DG}(r) = \frac{1}{Z_{DG}} \exp -\frac{1}{2} \frac{(\nu - \frac{\gamma\sqrt{1-\lambda}}{1-2\lambda})^2}{\lambda/(1-2\lambda)} \quad (24)$$

$$\text{with } Z_{DG} = \exp -\frac{1}{2} \left(\frac{\gamma^2}{1-2\lambda} + \log \frac{1-\lambda}{\lambda} \right) \quad (25)$$

B. Derivation of spike-count distribution via probability of common input

More easily, but somewhat less rigorously, the asymptotic spike count distribution can also be derived by first noting that the common input s and the normalized spike counts r are related via the change of variables from r to s via the function $r = L(s)$. The probability of observing r of the neurons spiking is the same as the probability of an input s

which has a probability $L(s)$ of producing a spike. Thus,

$$f(r) = \frac{\phi(L^{-1}(r))}{L'(L^{-1}(r))}. \quad (26)$$

This formula easily generalizes to arbitrary common input with distribution $f_s(s)$, and a noise distribution $P(X_i = 1|S_i = s_i) = L_i(s)$. In this case,

$$f(r) = \frac{f_s(L^{-1}(r))}{\prod_{i=1}^l L'_i(L_i^{-1}(r_i))}. \quad (27)$$

It is also easily generalizable to a population model with a finite number of pools, where the pairwise correlations within each pool, and across each pairs of pools are constant. Such a model is then fully described by the distributions over the vectors of spike-counts across each pool. For details, see [13].

C. Entropy rate of the DG

(We denote entropies by H and entropy rates by h , in slight departure from the notation in the letter, in which we used S and s). We note that in the entropy of the model can be split up into the entropy of the spike count, $H(K)$, and the conditional entropy given the spike count, i.e. $H(X|K)$, $H(X) = H(X|K) + H(K)$. The entropy of the the spike count $H(K)$, is bounded above by $\log(n)$, which is dominated by $H(X|k)$, which grows linearly. Hence, for large populations, we can calculate the entropy by considering only $H(X|k)$:

$$h_{DG} = \frac{1}{n} H(X) \quad (28)$$

$$= \frac{1}{n} \sum_k P(K = k) \log \binom{n}{k} \quad (29)$$

$$\approx \sum_k P(K = k) \left(-\frac{k}{n} \log \frac{k}{n} - \left(1 - \frac{k}{n}\right) \log \left(1 - \frac{k}{n}\right) \right) \quad (30)$$

$$\approx \int_0^1 f(r) [-r \log r - (1-r) \log(1-r)] dr \quad (31)$$

Alternatively, the entropy can be computed noting that it can be decomposed into the entropy of the model given the common input S , and the mutual information between common input and X , $MI(S : X)$:

$$H_{DG}(n) = - \sum_x p(x) \log(p(x)) \quad (32)$$

$$= - \int_{-\infty}^{\infty} \sum_x p(x, s) \log(p(x)) ds \quad (33)$$

$$= - \int_{-\infty}^{\infty} \sum_x p(x, s) \left(\log p(x|s) - \frac{\log p(x|s)}{p(x)} \right) ds \quad (34)$$

$$= - \int_{-\infty}^{\infty} p(s) \sum_x p(x|s) \log p(x|s) ds \quad (35)$$

$$+ \int_{-\infty}^{\infty} p(s) \sum_x p(x|s) \log \frac{p(x|s)}{p(x)} ds \quad (36)$$

$$=: H(X|S) + MI(S : X) \quad (37)$$

The conditional entropy $H(X|S)$ is given by

$$H(X|S) = n \int_{-\infty}^{\infty} p(s) [L(s) \log L(s) + L(-s) \log L(-s)] ds \quad (38)$$

The second term, $MI(S : X)$ is negligibly small for large n . This follows (e. g.) from the data-processing inequality:

$$MI(S : X) \leq MI(S : Z) \quad (39)$$

$$= H(Z) - H(Z|S) \quad (40)$$

$$= \frac{1}{2} (\log(\det(\Lambda)) - \log(\det((1 - \lambda)\mathbb{I}_n))) \quad (41)$$

$$= \frac{1}{2} \log \left(1 + \frac{\lambda n}{1 - \lambda} \right) \quad (42)$$

Therefore, for large n :

$$h_{DG} = \int_{-\infty}^{\infty} \phi_{\lambda}(s) [L(s) \log L(s) + L(-s) \log L(-s)] ds \quad (43)$$

By changing variables $L(s) = r$, one can see that the two derivations lead to the same result.

D. Heat capacity

Analogously to the entropy, the heat capacity and specific heat can easily be calculated from the spike-count distribution: We have that

$$HC = E(\log p(x))^2 - E(\log p(x))^2 \quad (44)$$

$$= \sum_x p(x) (\log p(x) - H(x))^2 \quad (45)$$

$$= \sum_k p(k) \left(\log p(k) - \log \binom{n}{k} - H \right)^2 \quad (46)$$

For large n , this can be approximated by

$$HC \approx \int_0^1 f(r) \left(\log f(r) - \log(n) - \log \binom{n}{nr} - hn \right)^2 \quad (47)$$

$$\approx \int_0^1 f(r) \log \binom{n}{nr}^2 + f(r) h^2 n^2 - 2f(r) \log \binom{n}{nr} h n dr \quad (48)$$

$$= n^2 \int_0^1 f(r) (\eta(r) - h)^2 dr \quad (49)$$

Hence, whenever this integral is non-zero, the heat capacity grows quadratically, and the specific heat linearly in n . As we pointed out, the integral will be non-zero for any homogeneous population model with interactions of at least order three.

III. ASYMPTOTIC PROPERTIES OF THE ISING MODEL

A. Entropy rate

For finite population sizes n , the spike-count distribution of the *MaxEnt* model is given by $P_{isi}(K = k) = \frac{1}{Z} \binom{n}{k} \exp(h_n k + J_n k^2)$. For large population sizes n , we chose to find the maximal entropy directly, without identifying the asymptotic scaling of the parameters h and J . (Alternatively, one could first identify the appropriate scaling of h and J , and then derive the entropy from the partition function, as, e.g., described in [16]. Note, however, that the 'usual' thermodynamic limit for the infinite range Ising model results in vanishing second order correlations, and is therefore not appropriate here.)

For large population sizes n , the entropy of the spike count $H(K)$ is negligibly small compared to the conditional entropy $H(X|K)$, and we find the MaxEnt distribution directly by finding the spike-count distribution P_{isi} which

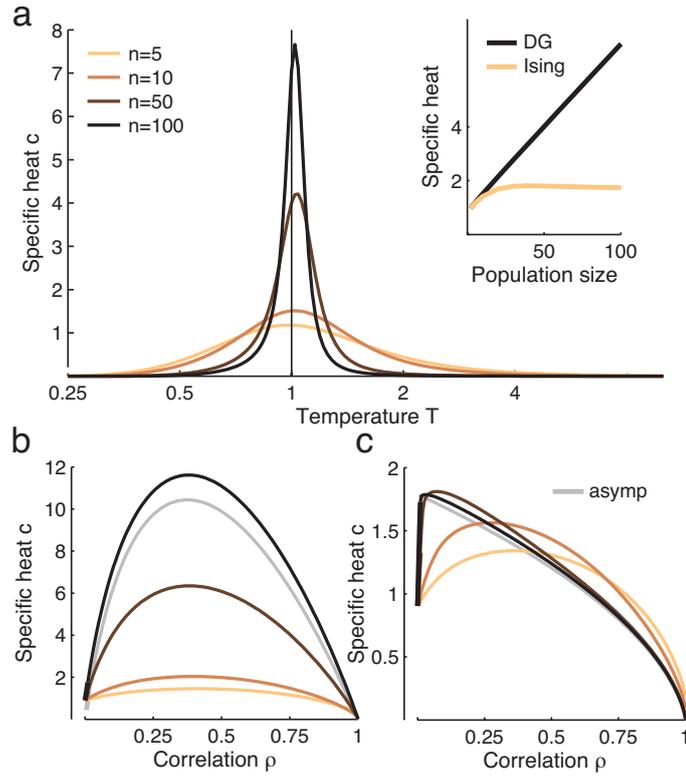


FIG. 2: Scaling of specific heat: **a)** Specific heat of the DG (for mean $\mu = 0.1$ and $\rho = 0.1$) diverges at $T = 1$. Inset: Specific heat of the DG at $T = 1$ grows linearly with population size. **b,c)** Specific heats for $\mu = 0.1$ and $T = 1$ vary with correlation ρ for DG (b) and Ising model (c) (gray: asymptotic heat, rescaled to be comparable to $n = 100$). For large n , the Ising model attains its maximum at values close to 0.

maximizes

$$H(X|K) = \frac{1}{n} \sum_k P_{isi}(K = k) \log \binom{n}{k}. \quad (50)$$

This is really the linear program

$$\max_P H(X|K) = \frac{1}{n} \sum_k P_k \log \binom{n}{k} \quad (51)$$

$$\text{subject to } P_k \geq 0 \forall k \quad (52)$$

$$\sum_k P_k = 1 \quad (53)$$

$$\frac{1}{n} \sum_k P_k k = \mu \quad (54)$$

$$\frac{1}{n^2} \sum_k P_k k^2 = \mu^2 n^2 + \sigma n(n-1) + \mu(1-\mu)n. \quad (55)$$

$$(56)$$

In the limit of $n \rightarrow \infty$, we can write it as

$$\max_f H(X|K) = - \int_0^1 f(r)r \log(r) + (1-r) \log(1-r) dr \quad (57)$$

$$\text{subject to } f(r) \geq 0 \quad \forall r \in [0, 1] \quad (58)$$

$$\int f(r) dr = 1 \quad (59)$$

$$\int f(r)r dr = \mu \quad (60)$$

$$\int f(r)r^2 dr = \mu^2 + \sigma \quad (61)$$

$$(62)$$

By finding the function f that maximizes it, we can find the asymptotic entropy of the maximum-entropy distribution without having to use the exponential form of the MaxEnt distribution, and in particular, without identifying the parameters h and J . We make the informed guess that the functional form of P which maximizes is a mixture of two delta-distributions, $f_{isi}(r) = p_1\delta(r - r_1) + p_2\delta(r - r_2)$, where $r_2 = 1 - r_1$. The reason for this is that to get an asymptotic distribution with non-vanishing correlations, we need at least two different 'states' of the population, corresponding to the physical system being in a ferromagnetic, rather than paramagnetic phase. Note that the 'usual' scaling of J as $1/n$ and h as constant yields a paramagnetic solution with vanishing correlations.

The locations $r_{1,2}$ and probabilities $p_{1,2}$ are chosen such that the mean and variance are as desired, i.e. by

$$\mu = p_1 r_1 + p_2 r_2 \quad (63)$$

$$\sigma + \mu^2 = p_1 r_1^2 + p_2 r_2^2 \quad (64)$$

$$r_{1,2} = \frac{1}{2} \pm \sqrt{1/4 - \mu + \mu^2 + \sigma} \quad (65)$$

$$p_2 = \frac{\mu + r_2 - 1}{2r_2 - 1} \quad (66)$$

$$p_1 = 1 - p_2. \quad (67)$$

The entropy rate is given by

$$h_{isi} = \lim_{n \rightarrow \infty} \left(\frac{p_1}{n} \log \binom{n}{nr_1} + \frac{p_2}{n} \log \binom{n}{nr_2} \right) \quad (68)$$

$$= p_1 \eta_2(r_1) + p_2 \eta_2(r_2) \quad (69)$$

$$\text{where } \eta_2(r) = -r \log_2(r) - (1-r) \log_2(1-r) \quad (70)$$

Summing up, the entropy is given by:

$$r = \frac{1}{2} \pm \sqrt{1/4 - \mu + \mu^2 + \sigma} \quad (71)$$

$$= \frac{1}{2} \pm \sqrt{1/4 + \mu(1-\mu)(1-\rho)} \quad (72)$$

$$\eta_2 = -r \log_2(r) - (1-r) \log_2(1-r) \quad (73)$$

Thus, we can calculate the asymptotic entropy rate of the Ising model without having to identify the parameters h and J explicitly. Also, we see that an Ising model with rate μ_0 and no correlations has the same entropy as one rate μ_1 with pairwise covariance σ provided that $\mu_0^2 - \mu_0 = \mu_1^2 - \mu_1 + \sigma$.

We are yet to show that the form of f_{isi} that we guessed above is indeed the solution to the linear program. In general, one can prove the optimality [10] of a feasible solution x^* for the linear program

$$\min v^\top x \quad (74)$$

$$\text{subject to } x \geq 0 \quad (75)$$

$$Ax = b \quad (76)$$

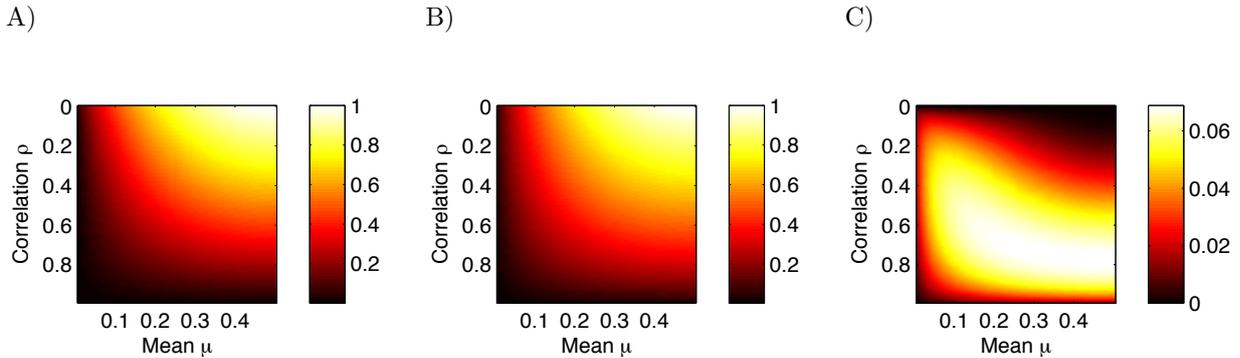


FIG. 3: Asymptotic Entropy of the DG and MaxEnt model **A**) Asymptotic entropy rate of the DG-model as a function of mean μ and correlation coefficient ρ **B**) Asymptotic entropy rate of the corresponding MaxEnt model **C**) Difference in entropy, i.e. KL-divergence, between the two models

by finding a solution λ^* to the dual program

$$\max \lambda^\top b \quad (77)$$

$$A^\top \lambda \leq v \quad (78)$$

which has the same result, i.e. $v^\top x^* = \lambda^{*\top} b$. Rewriting the MaxEnt problem as minimizing the negative entropy, we need to find a vector λ which satisfies

$$\text{a) } \lambda_1 + \lambda_2 \mu + \lambda_3 (\mu^2 + \sigma) = -h_{isi} \quad (79)$$

$$\text{b) } \lambda_1 + \lambda_2 s + \lambda_3 s^2 \leq s \log(s) + (1-s) \log(1-s). \quad \forall s \in [0, 1], \quad (80)$$

We take

$$\lambda_1 = -h_{isi} + \lambda_2(r^2 - r) \quad (81)$$

$$\lambda_2 = \log\left(\frac{r}{1-r}\right) / (1-2r) \quad (82)$$

$$\lambda_3 = -\lambda_2. \quad (83)$$

(The values of λ are obtained by requiring that the equality a) holds, and that inequality b) holds with equality at the point $s = \mu - \mu^2 - c$, and that the derivatives of the lhs and rhs of equation b) also match at this point).

Noting that $\mu - \mu^2 - c = r - r^2$, condition a) holds:

$$\lambda_1 + \lambda_2 \mu + \lambda_3 (\mu^2 + \sigma) = \quad (84)$$

$$= -h_{isi} + \lambda_2(r^2 - r) + \lambda_2(\mu - \mu^2 - \sigma) \quad (85)$$

$$= -h_{isi} \quad (86)$$

For condition b) to hold, we check that r is a local minimum of the function $\Delta(s) = rhs(s) - lhs(s)$. As $\Delta(r) = 0$, and that it does not have any other stationary points for $s \leq 1/2$, we get that $\Delta(s) \geq 0 \forall s \in (0, 1)$.

B. Derivation of large- n spike count distribution

We want to identify the parameters h_n and J_n of the Ising model in the parameterization $P(K = k) = \binom{n}{k} \exp(h_n k + J_n k^2)$. For large n , the distribution of normalized spike counts R is given by

$$f(r) = \frac{1}{Z} \exp(n(\eta_e + h_n r + n J_n^2 r^2)) \quad (87)$$

We know that the spike-count distribution asymptotically collapses to two delta-peaks at locations $r_{1,2}$ with weights $p_{1,2}$. Thus, we know that

$$\log(f(r_1)) - \log(f(r_2)) \rightarrow \log(p_1) - \log(p_2) \quad (88)$$

$$\text{also } \log(f(r_1)) - \log(f(r_2)) = n(h_n + nJ_n)(1 - 2r_1) \quad (89)$$

$$\text{and thus } n(h_n + nJ_n)(1 - 2r_1) \rightarrow \log(p_1) - \log(p_2) \quad (90)$$

Thus, we need $J_n = J/n$ and h of order 1, and for large n , $(h_n + nJ_n) = \frac{1}{n} \frac{\log p_1 - \log p_2}{r_2 - r_1} =: \frac{\alpha}{n}$. We also know that in the limit, we need $h_n + r_1 n J_n = \log(r_1) - \log(r_2)$ to ensure that first local maximum of $f(r)$ is at r_1 .

To satisfy these constraints, we set

$$J_n = \frac{1}{n} \frac{\log r_2 - \log r_1}{r_2 - r_1} =: J/n \quad (91)$$

$$h_n = \alpha/n - J \quad (92)$$

Summing up, we get

$$f_{isi}(r) = \frac{1}{Z_{isi}} \exp(\alpha r + n(\eta_e(r) + J(r^2 - r))) \quad (93)$$

$$Z_{isi} = n \sqrt{\frac{2\pi}{n\left(\frac{1}{r_1 r_2} - 2J\right)}} (\exp(\alpha r_1) + \exp(\alpha r_2)) \exp(n(\eta_e(r) + J(r_1^2 - r_1))), \quad (94)$$

where the partition function Z_{isi} can be derived using a saddle-point approximation around r_1 and r_2 . This large- n approximation over-estimates the spike-count variance, and thus the pairwise correlations slightly (by the variance σ_n defined in the next section). Therefore, its estimates for the sparsity are biased downwards.

C. Heat capacity

We want to calculate the asymptotic scaling of the specific heat of the Ising model. We first note that the diverging part of the specific heat, which we calculated for the DG , is zero for the (homogeneous) Ising model:

$$c_{isi} = n \int f(r) (\eta_2(r)^2 - s_\beta^2) dr \quad (95)$$

$$= p_1 (\eta_2^2(r_1) - \eta_2^2(r_1)) + p_2 (\eta_2^2(r_2) - \eta_2^2(r_2)) \quad (96)$$

$$= 0 + 0 \quad (97)$$

Thus, we know that the specific heat remains finite, and all terms corresponding to the linear terms cancel out to give zero. (In general, this integral could be non-zero for a pair of delta-peaks, but not if they are symmetric around 0.5, as in this case, η_2 evaluate at the delta-peaks will be exactly the entropy, i.e. the term inside the square will be 0.)

For large n , $f(r)$ can be approximated using a mixture of two Gaussians with means r_1 and r_2 , variances $\sigma_n = \left(n\left(\frac{1}{r_1 r_2} - 2J\right)\right)^{-1}$, and relative heights p_1 and p_2 , $f_n(r) = p_1 \mathcal{N}(r, r_1, \sigma_n) + p_2 \mathcal{N}(r, r_2, \sigma_n)$.

The specific heat c_{isi} is given by

$$c_{isi} = \frac{1}{n} \text{Var}(\log_2 p(x)) \quad (98)$$

$$= \frac{1}{n} \left(E_K (h_n k + J k^2)^2 - E_K (h_n k + J k^2)^2 \right) \quad (99)$$

$$= n \int f_n(r) (h_n r + n J_n r^2)^2 - \left(\int f_n(r) (h_n r + n J_n r^2) dr \right)^2 \quad (100)$$

$$= n (h_n^2 (E(r^2) - E(r)^2) + n^2 J_n^2 (E(r^4) - E(r^2)^2) + 2 J_n h_n n (E(r^3) - E(r^2)E(r))) \quad (101)$$

$$:= n (h_n^2 C_2 + n^2 J_n^2 C_4 + 2 J_n h_n n C_3) \quad (102)$$

$$\approx n ((\alpha/n - J)^2 C_2 + J^2 C_4 + 2 J (\alpha/n - J) C_3) \quad (103)$$

$$= n J^2 (C_4 + C_2 - 2 C_3) + 2 \alpha J (C_3 - C_2) + \frac{\alpha^2}{n} C_2 \quad (104)$$

where the expectations are with respect to f_n .

Using

$$E(r^4) = p_1 r_1^4 + p_2 r_2^4 + 6 \sigma_n (\mu^2 + \sigma) + 3 \sigma_n^2 \quad (105)$$

$$E(r^3) = p_1 r_1^3 + p_2 r_2^3 + 3 \sigma \mu \quad (106)$$

$$E(r^2) = \sigma + p_1 r_1^2 + p_2 r_2^2 \quad (107)$$

we get that

$$C_4 + C_2 - 2 C_3 = \sigma_n (4(\sigma + \mu^2) - 4\mu + 1) \quad (108)$$

$$C_3 - C_2 = \sigma (2\mu - 1) \quad (109)$$

This firstly confirms that the linear term in c_{isi} drops out, and that (ignoring terms which are of order $1/n$ or $1/n^2$, the specific heat of the Ising model at unit-temperature is given by

$$c_{isi} = n J^2 \sigma_n (4(\sigma + \mu^2) - 4\mu + 1) \quad (110)$$

$$= \frac{J^2 (4(\sigma + \mu^2) - 4\mu + 1)}{\frac{1}{r_1 r_2} - 2J} \quad (111)$$

$$= \frac{J^2 (4(\sigma + \mu^2) - 4\mu + 1)}{\frac{1}{r_1 r_2} - 2J} \quad (112)$$

Multiplying by $\log_2^2(e)$ to rescale to log base 2 yields the result.

IV. EXAMPLE OF A SYSTEM WITH BOUNDED (OR ZERO) CORRELATIONS AND A NON-EXTENSIVE ENTROPY

In case of independent neurons the entropy of all neurons adds up linearly leading to the maximal possible increase in entropy. Any kind of redundancy reduces the total entropy but in many cases entropy is still extensive in the number of neurons. Therefore, it is not obvious under which conditions a sublinear scaling of entropy in the limit of large n can be achieved with bounded second-order correlations. The simplest example may be the "grandmother cell" code where each neuron in a population encodes for a different stimulus such that only one neuron is active at any given time. Assuming that the probability of being activated is the same for all neurons, the resulting entropy is $\log(n)$. Thus the entropy rate $\log_2(n)/n$ goes to zero for large n .

In the grandmother cell code example, however, not only the entropy rate but the probability of firing $\mu = 1/n$ and hence the entropy per neuron $(\log_2 n + (n-1) \log_2(n/(n-1)))/n$ goes to zero for large n as well. In other words, the sublinear growth of the total entropy is not due to an increasing amount of redundancy but due to the vanishing entropy of the individual neurons.

In the following, we will show that it is possible to construct a code for which the entropy of each neuron is fixed and the joint entropy still scales sublinearly even though the pairwise correlations vanish. Like in the grandmother cell

code, we will again assume that n neurons will encode for n different stimuli only. In contrast to the grandmother cell code, however, we will now assume that each neuron has a tuning curve which is given by a random, binary vector: For randomly chosen $n/2$ out of the possible stimulus values, it deterministically spikes, whereas for the other $n/2$ values, it deterministically stays silent. In this setup, each neuron has a firing rate close to 0.5, and pairwise correlations vanish for large n . However, the entropy is not extensive, but still given by $\log_2(n)$, as there is a deterministic mapping of population-states to the n stimulus values.

Finally, we remark that the example can be modified to have exactly zero correlation across all pairs of neurons when using non-random tuning curves and $(n + 1)$ stimulus values. For $n = 2^m - 1$ it is always possible to construct an $(n + 1) \times (n + 1)$ Hadamard matrix Q for which $Q \in \{-1, 1\}$ and $QQ' = (n + 1)I$. One row of Q consists only of ones while for all other rows exactly half of the entries are one and the other half are minus one. Let F denote the $n \times (n + 1)$ matrix which is obtained from Q by leaving out the row with ones only. If $e_k \in R^{(n + 1)}$ is the k -th standard basis vector whose k -th component is one and zero otherwise, then we can describe the activation of the neurons for any stimulus value k by Fe_k . If we again assume that the stimuli are distributed uniformly, the correlation between all pairs of neurons is zero due to the orthogonality of the Hadamard matrix. Thus we have constructed an example for which the firing probability of each neuron is exactly 0.5 (and hence the entropy of each neuron 1bit), all pairwise correlations are exactly zero, and the entropy scaling is given by $\log_2(n + 1)$.

Acknowledgments

JHM and MB were supported by the Bernstein Prize to MB (BMBF; FKZ: 01GQ0601), the Max Planck Society and the Werner Reichardt Centre for Integrative Neuroscience. JHM was additionally supported by a Marie Curier Intra European Fellowship from the European Commission. We thank S Gerwinn and F Sinz for very useful discussions in particular regarding the example in section 4, Eran Mukamel, Sebastian Gerwinn and Peter Latham for comments on the manuscript, and S. Liebe for help with figures.

-
- [1] Shun-Ichi Amari, Hiroyuki Nakahara, Si Wu, and Yutaka Sakai. Synchronous firing and higher-order interactions in neuron pool. *Neural Comp*, 15(1):127–142, 2003.
 - [2] Andrea K. Barreiro, Julijana Gjorgjieva, Fred Rieke, and Eric Shea-Brown. When are feedforward microcircuits well-modeled by maximum entropy methods. *arXiv:1011.2797v2 [q-bio.NC]*, 11/2010.
 - [3] M. Bethge and P. Berens. Near-maximum entropy models for binary neural representations of natural images. In *Neural Information Processing Systems 20*, volume 21, 2007.
 - [4] S M Bohte, H Spekreijse, and P R Roelfsema. The effects of pair-wise and higher order correlations on the firing rate of a post-synaptic neuron. *Neural Comput*, 12(1):153–179, 2000.
 - [5] R. Brette. Generation of correlated spike trains. *Neural Computation*, 21(1):188–215, 2009.
 - [6] Y. Burak, S. Lewallen, and H. Sompolinsky. Stimulus-dependent correlations in threshold-crossing spiking neurons. *Neural computation*, 21(8):2269–2308, 2009.
 - [7] Simona Cocco, Stanislas Leibler, and Remi Monasson. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. *Proc Natl Acad Sci U S A*, 106(33):14058–14062, 2009 Aug 18.
 - [8] D. R. Cox and N. Wermuth. On some models for multivariate binary variables parallel in complexity with the multivariate gaussian distribution. *Biometrika*, 89:462–469, 2002.
 - [9] Jessy D Dorn and Dario L Ringach. Estimating membrane voltage correlations from extracellular spike trains. *J Neurophysiol*, 89(4):2271–2278, 2003 Apr.
 - [10] W. Karush. Minima of functions of several variables with inequalities as side constraints. Master’s thesis, University of Chicago, 1939.
 - [11] M. Krumin and S. Shoham. Generation of Spike Trains with Controlled Auto-and Cross-Correlation Functions. *Neural Computation*, pages 1–23, 2009.
 - [12] Jakob H Macke, Philipp Berens, Alexander S Ecker, Andreas S Tolias, and Matthias Bethge. Generating spike trains with specified correlation coefficients. *Neural Comp*, 2009.
 - [13] Jakob H Macke, Manfred Opper, and Matthias Bethge. The effect of pairwise correlations on global population statistics. Technical report, Max Planck Institute for Biological Cybernetics, 2009.
 - [14] O Marre, S El Boustani, Y Fregnac, and A Destexhe. Prediction of spatiotemporal patterns of neural activity from pairwise correlations. *Phys Rev Lett*, 102(13):138101, 2009 Apr 3.
 - [15] Fernando Montani, Robin A A Ince, Riccardo Senatore, Ehsan Arabzadeh, Mathew E Diamond, and Stefano Panzeri. The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. *Philos Transact A Math Phys Eng Sci*, 367(1901):3297–3310, 2009 Aug 28.
 - [16] H. Nishimori. *Statistical physics of spin glasses and information processing*. Oxford University Press New York, 2001.

- [17] IE Ohiorhenuan and JD Victor. Information-geometric measure of 3-neuron firing patterns characterizes scale-dependence in cortical networks. *J Comput Neurosci*, 30:125–141, 2011.
- [18] K. Pearson. On a new method of determining correlation between a measured character a, and a character b, of which only the percentage of cases wherein b exceeds (or falls short of) a given intensity is recorded for each grade of a. *Biometrika*, 7:96–105, 1909.
- [19] Yasser Roudi, Erik Aurell, and John A Hertz. Statistical physics of pairwise probability models. *Front Comput Neurosci*, 3:22, 2009.
- [20] Yasser Roudi, Sheila Nirenberg, and Peter E Latham. Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. *PLoS Comput Biol*, 5(5):e1000380, 2009.
- [21] Yasser Roudi, Joanna Tyrcha, and John Hertz. Ising model for neural data: model quality and approximate methods for extracting functional connectivity. *Phys Rev E Stat Nonlin Soft Matter Phys*, 79(5 Pt 1):051915, 2009.
- [22] Elad Schneidman, Michael J 2nd Berry, Ronen Segev, and William Bialek. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440(7087):1007–12, 2006.
- [23] Jonathon Shlens, Greg D Field, Jeffrey L Gauthier, Martin Greschner, Alexander Sher, Alan M Litke, and E J Chichilnisky. The structure of large-scale synchronized firing in primate retina. *Journal of Neuroscience*, 29(15):5022–5031, 2009.
- [24] Jonathon Shlens, Greg D Field, Jeffrey L Gauthier, Matthew I Grivich, Dumitru Petrusca, Alexander Sher, Alan M Litke, and E J Chichilnisky. The structure of multi-neuron firing patterns in primate retina. *J Neurosci*, 26(32):8254–66, 2006.
- [25] Aonan Tang, David Jackson, Jon Hobbs, Wei Chen, Jodi L Smith, Hema Patel, Anita Prieto, Dumitru Petrusca, Matthew I Grivich, Alexander Sher, Pawel Hottowy, Wladyslaw Dabrowski, Alan M Litke, and John M Beggs. A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. *Journal of Neuroscience*, 28(2):505–518, 2008.
- [26] T. Tchumatchenko, A. Malyshev, T. Geisel, M. Volgushev, and F. Wolf. Correlations and synchrony in threshold neuron models. *Physical review letters*, 104(5):58102, 2010.