# Statistical mechanics for real biological networks

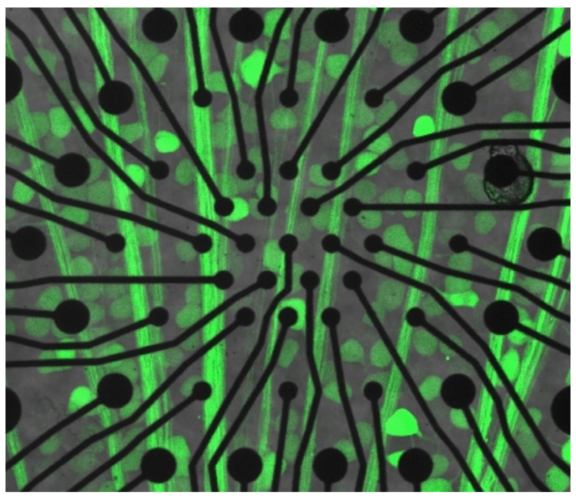

William Bialek

Joseph Henry Laboratories of Physics, and
Lewis-Sigler Institute for Integrative Genomics
Princeton University

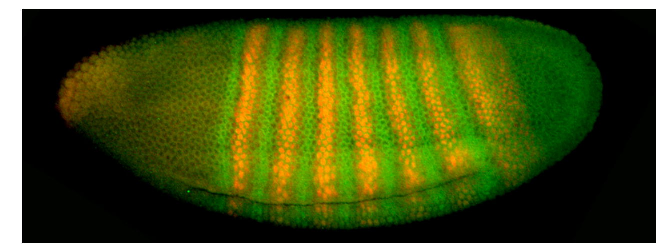
Initiative for the Theoretical Sciences
The Graduate Center
City University of New York

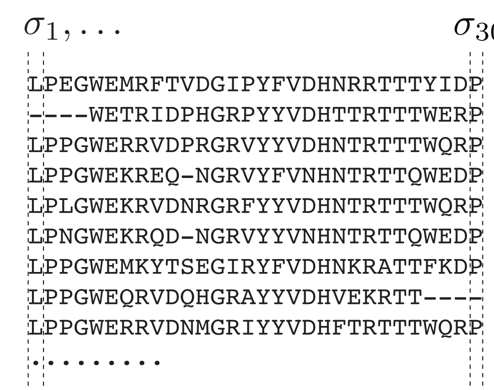

**Colloquium, Santa Fe Institute**
**Thursday, 27 October, 2011**

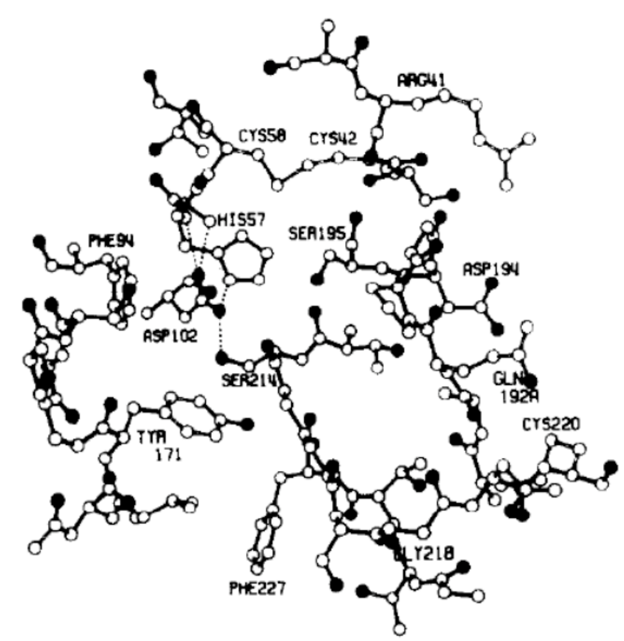
# Many of the phenomena of life emerge from interactions among large numbers of simpler elements: A network



| state of one element in the network | state of the whole network | how do we get from one state to the next? | which states are sampled in real life? |
|---|---|---|---|
| spike/silence from one neuron | a percept, memory, motor plan or thought | ion channels and synapses | coding capacity, multiple attractors, ... ? |
| expression level of one gene | fate of a cell during embryonic development | diffusion, binding, “walking,” ... | how many fates? modules vs. emergence? |
| choice of amino acid at one site in a protein | sequence: structure & function | evolution | how large is the repertoire? |
| flight direction and speed of one bird in a flock | coherent migration | sensorimotor response + Newton | propagation of information, responsiveness, ... ? |
| $\sigma_{\rm n}$ | $\{\sigma_{\rm n}\}$ | $\{\sigma_{\rm n}\}_t \rightarrow \{\sigma_{\rm n}\}_{t+\Delta t}$ | $P\left(\{\sigma_{\rm n}\}\right)$ |





# There is a long history of physicists trying to use ideas from statistical mechanics to approach these systems ...

*Proc. Natl. Acad. Sci. USA*
Vol. 79, pp. 2554–2558, April 1982
Biophysics

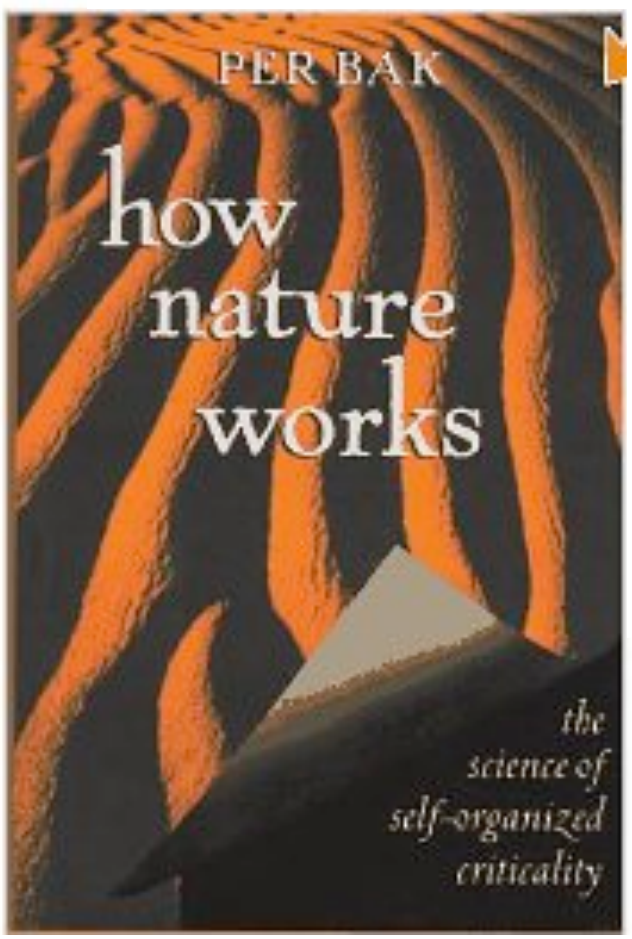
**Neural networks and physical systems with emergent collective computational abilities**

(associative memory/parallel processing/categorization/content-addressable memory/fail-soft devices)

J. J. HOPFIELD

approximate dynamics → Lyapunov function ≈ "energy landscape" on space of states

stable states (stored memories, answers to computations, ...) are energy minima




If there is a statistical mechanics of network states, there is a phase diagram as networks get large ...
Where are we in this space?
On a critical surface, perhaps?

But how do we connect these ideas to real data?

# Once again, what did we want to know about the system?

How do we get from one state to another?

$$\{\sigma_{\rm n}\}_t \to \{\sigma_{\rm n}\}_{t+\Delta t}$$

Which states of the whole system are sampled in real life?

$$P\left(\{\sigma_{\rm n}\}\right)$$

State space is too large to answer these (and other) questions "directly" by experiment.

You can't measure all the states, but you can measure averages, correlations, ... .

Could we build the minimally structured model that is consistent with these measurements?

"Minimally structured" = maximum entropy, and this connects the real data directly to statistical mechanics ideas (!).

**Network information and connected correlations.**
E Schneidman, S Still, MJ Berry II & W Bialek, *Phys Rev Lett* 91, 238701 (2003); arXiv:physics/0307072 (2003).

**Weak pairwise correlations imply strongly correlated network states in a neural population.**
E Schneidman, MJ Berry II, R Segev & W Bialek, *Nature* 440, 1007-1012 (2006); arXiv:q-bio.NC/0512013 (2005).

**Ising models for networks of real neurons.**
G Tkacik, E Schneidman, MJ Berry II & W Bialek, arXiv:q-bio.NC/0611072 (2006).

**Rediscovering the power of pairwise interactions.**
W Bialek & R Ranganathan, arXiv:0712.4397 (2007).

**Statistical mechanics of letters in words.**
GJ Stephens & W Bialek, *Phys Rev E* 81, 066119 (2010); arXiv:0801,0253 [q-bio.NC] (2008).

**Spin glass models for networks of real neurons.**
G Tkacik, E Schneidman, MJ Berry II & W Bialek, arXiv:0912.5409 [q-bio.NC] (2009).

**Maximum entropy models for antibody diversity.**
T Mora, AM Walczak, W Bialek & CG Callan, Jr, *Proc Nat'l Acad Sci (USA)* 107, 5405-5410 (2010); arXiv:0912.5175 (2009).

**Are biological systems poised at criticality?**
T Mora & W Bialek, *J Stat Phys* 144, 268-302 (2011); arXiv:1012.2242 [q-bio.QM] (2010).

**When are correlations strong?**
F Azhar & W Bialek, arXiv:1012.5987 [q-bio.NC] (2010).

**Statistical mechanics for a natural flock of birds**
W Bialek, A Cavagna, I Giardina, T Mora, E Silverstri, M Viale & AM Walczak, arXiv:1107.0604 [physics.bio-ph] (2011).

**and, most importantly, work in progress.**

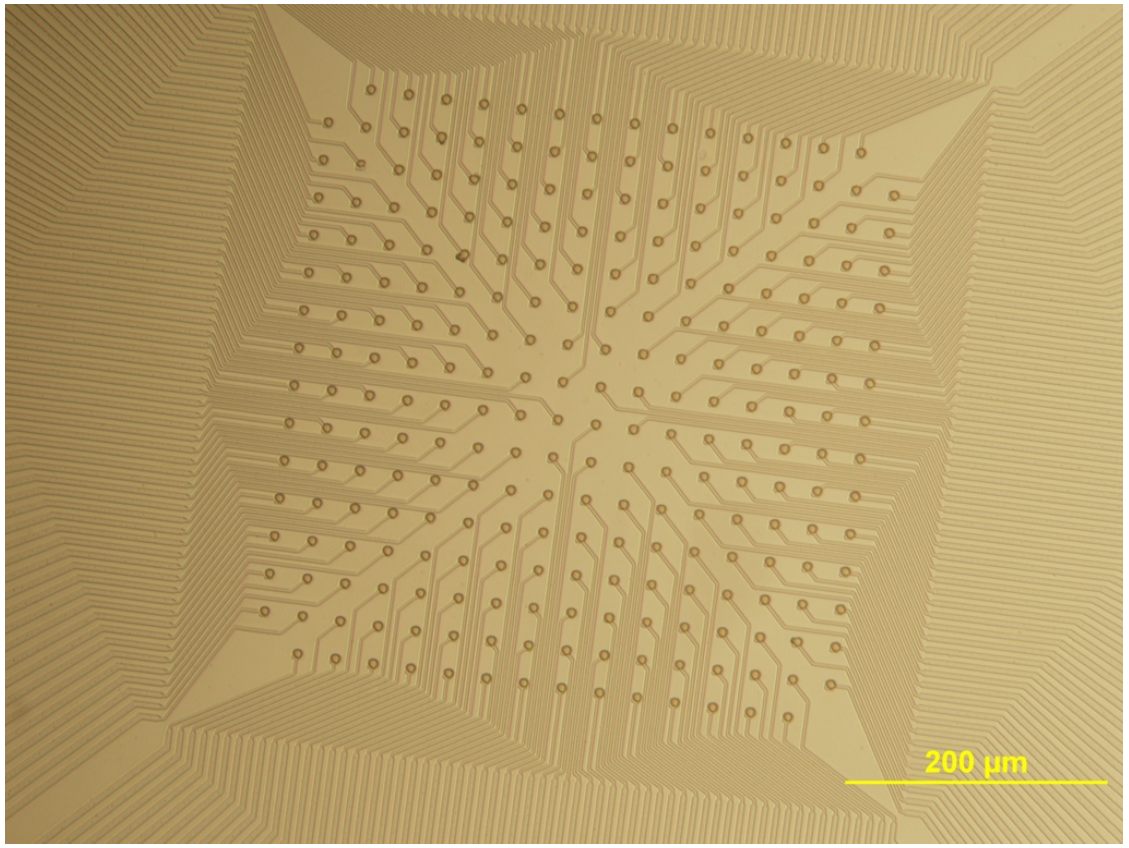
1. Basics of maximum entropy construction

2. The case of neurons in the retina

3. A bit about proteins, birds, ...

4. Perspective

(a few words about the retinal experiments)




D Amodei, O Marre & MJ Berry II, in preparation (2011).

Let's define various function of the state of the system,

$$f_1(\{\sigma_{\rm n}\}), f_2(\{\sigma_{\rm n}\}), \cdots, f_K(\{\sigma_{\rm n}\}),$$

and assume that experiments can tell us the averages of these functions:

$$\langle f_1(\{\sigma_{\rm n}\})\rangle, \langle f_2(\{\sigma_{\rm n}\})\rangle, \cdots, \langle f_K(\{\sigma_{\rm n}\})\rangle.$$

What is the least structured distribution $P(\{\sigma_{\rm n}\})$ that can reproduce these measured averages?

This tells us the <u>form</u> of the distribution.

$$P(\{\sigma_{\rm n}\}) = \frac{1}{Z(\{g_\mu\})} \exp\left[-\sum_{\mu=1}^{K} g_\mu f_\mu(\{\sigma_{\rm n}\})\right]$$

Matching expectation values = maximum likelihood inference of parameters.

Still must adjust the coupling constants $\{g_\mu\}$ to match the measured expectation values.

Reminder: Suppose this is a physical system, and there is some energy for each state, $E(\{\sigma_{\rm n}\})$

Thermal equilibrium is described by a distribution that is as random as possible (maximum entropy) while reproducing the observed average energy:

$$P(\{\sigma_{\rm n}\}) = \frac{1}{Z} \exp\left[-\beta E(\{\sigma_{\rm n}\})\right]$$

In this view, the temperature $T = 1/(k_B \beta)$ is just a parameter we adjust to reproduce $\langle E(\{\sigma_{\rm n}\})\rangle$.

We can think of the maximum entropy construction as defining an effective energy for every state,

$$E(\{\sigma_{\rm n}\}) = \sum_{\mu=1}^{K} g_\mu f_\mu(\{\sigma_{\rm n}\}), \qquad \text{with}\, k_B T = 1.$$

This is an exact mapping, not an analogy.

---

## Examples for neurons:

$\{\langle f_\mu \rangle\} =$

probability of spike vs. silence $\Rightarrow E = \sum_{\rm n} h_{\rm n}\sigma_{\rm n}$

spike probability + pairwise correlations $\Rightarrow E = \sum_{\rm n} h_{\rm n}\sigma_{\rm n} + \frac{1}{2}\sum_{\rm nm} J_{\rm nm}\sigma_{\rm n}\sigma_{\rm m}$

probability that M cells spike together $\Rightarrow E = V\left(\sum_{\rm n} \sigma_{\rm n}\right)$ this case we can do analytically

# maximum entropy model consistent with probability of M cells spiking together

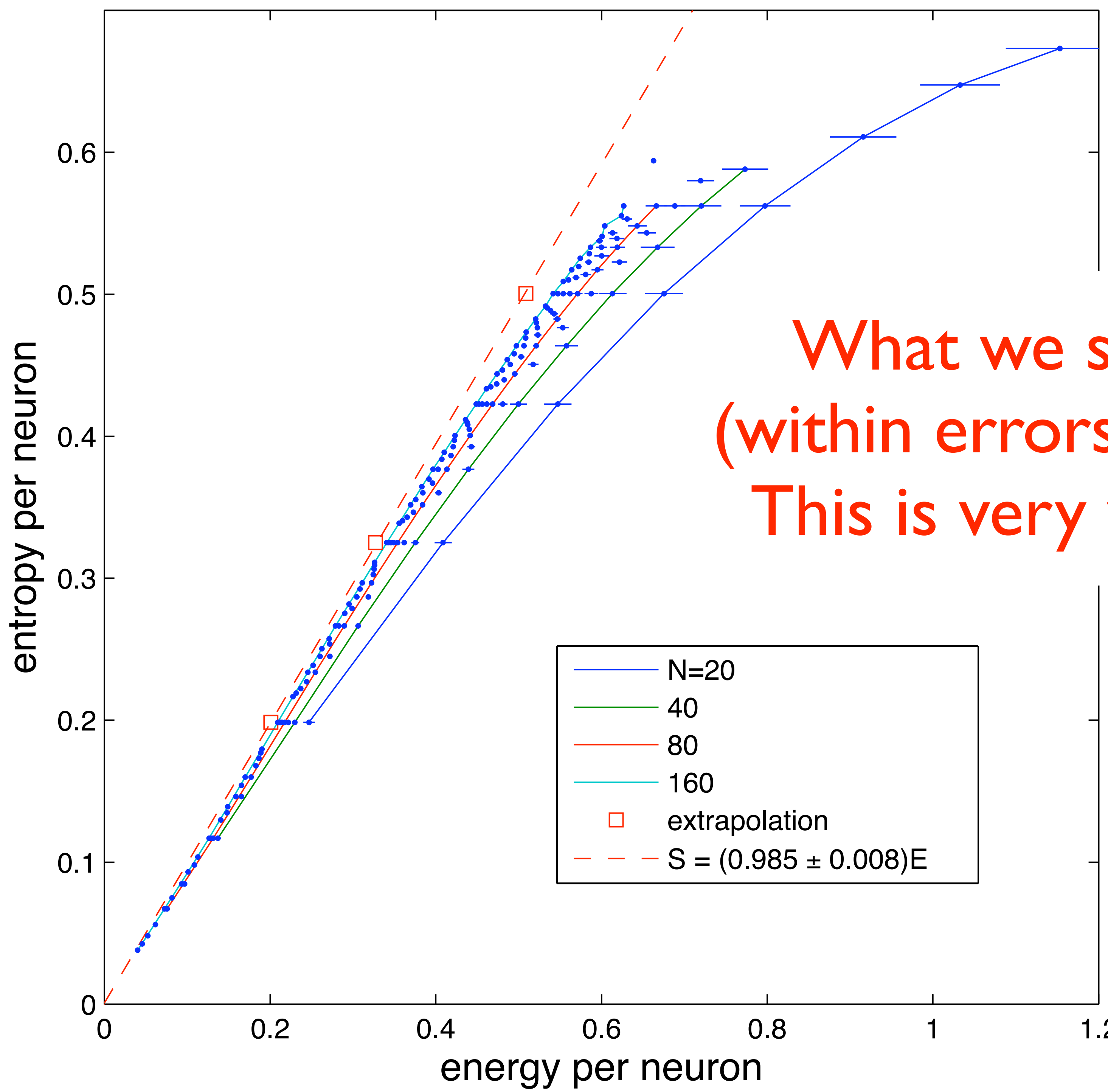
$$\Rightarrow E = V\left(\sum_{n} \sigma_{n}\right)$$

Find this global "potential" for multiple subgroups of N neurons.

Lots of states with the same energy ... count them to get the entropy.

For large N we expect entropy and energy both proportional to N.

Plot of S/N vs. E/N contains all the "thermodynamics" of the system.



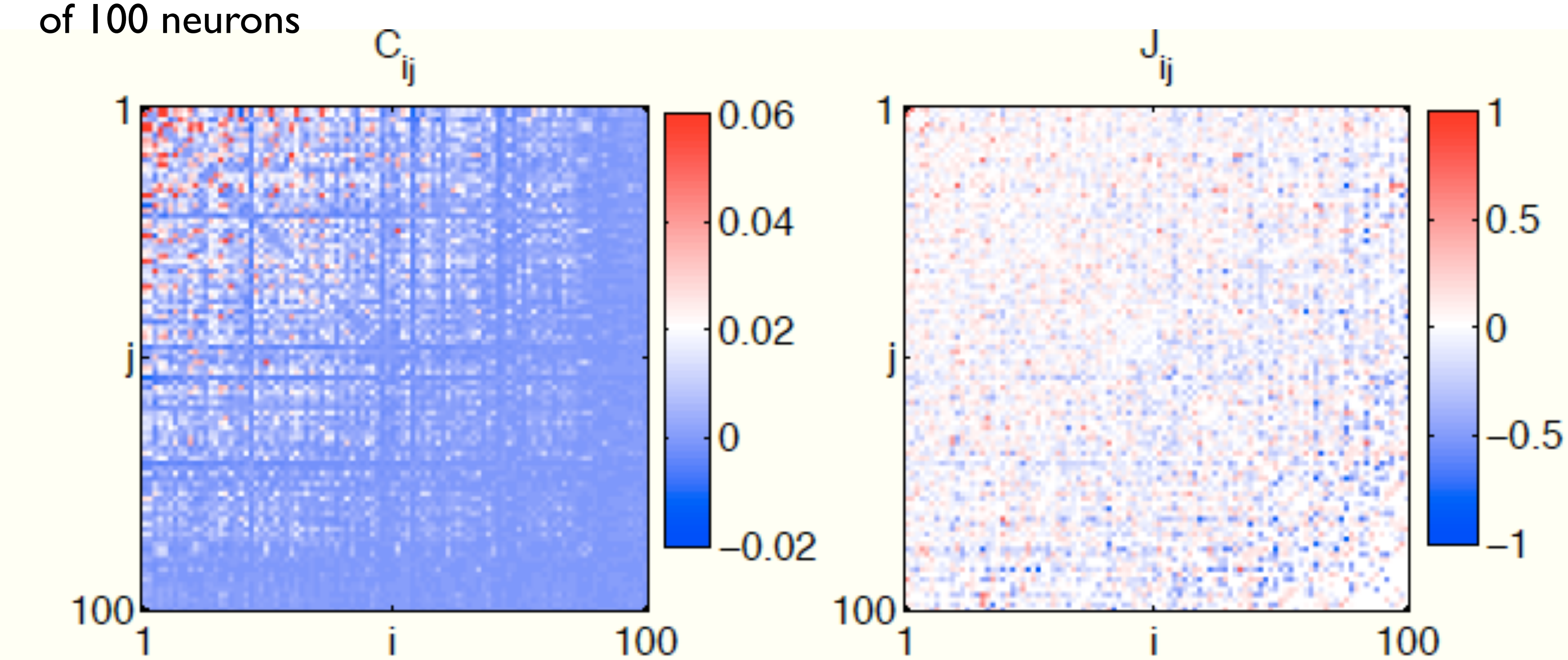

What we see is (within errors) S = E. This is very weird.

The real problem: maximum entropy model consistent with mean spike probabilities, pairwise correlations, and probabilities of M cells spiking together.

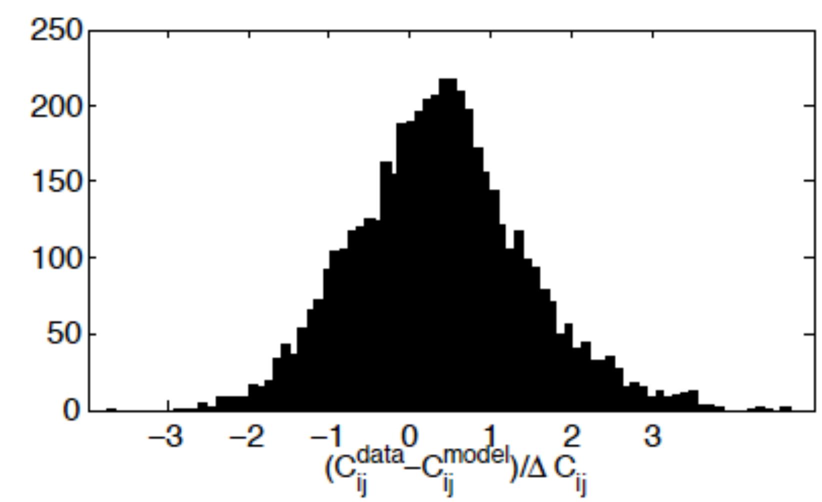
$$E = \sum_{\mathrm{n}} h_{\mathrm{n}} \sigma_{\mathrm{n}} + \frac{1}{2} \sum_{\mathrm{nm}} J_{\mathrm{nm}} \sigma_{\mathrm{n}} \sigma_{\mathrm{m}} + V \left( \sum_{\mathrm{n}} \sigma_{\mathrm{n}} \right)$$

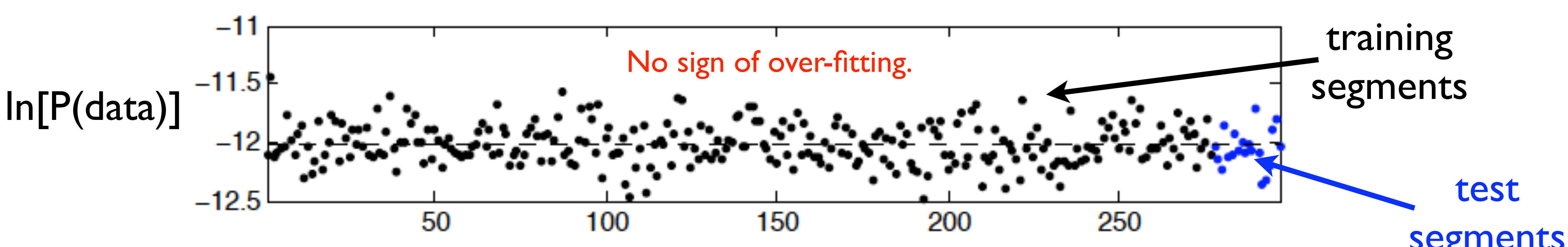
There are lots of parameters, but we can find them all from ~1 hour of data. This is a hard "inverse statistical mechanics" problem.

one example of 100 neurons

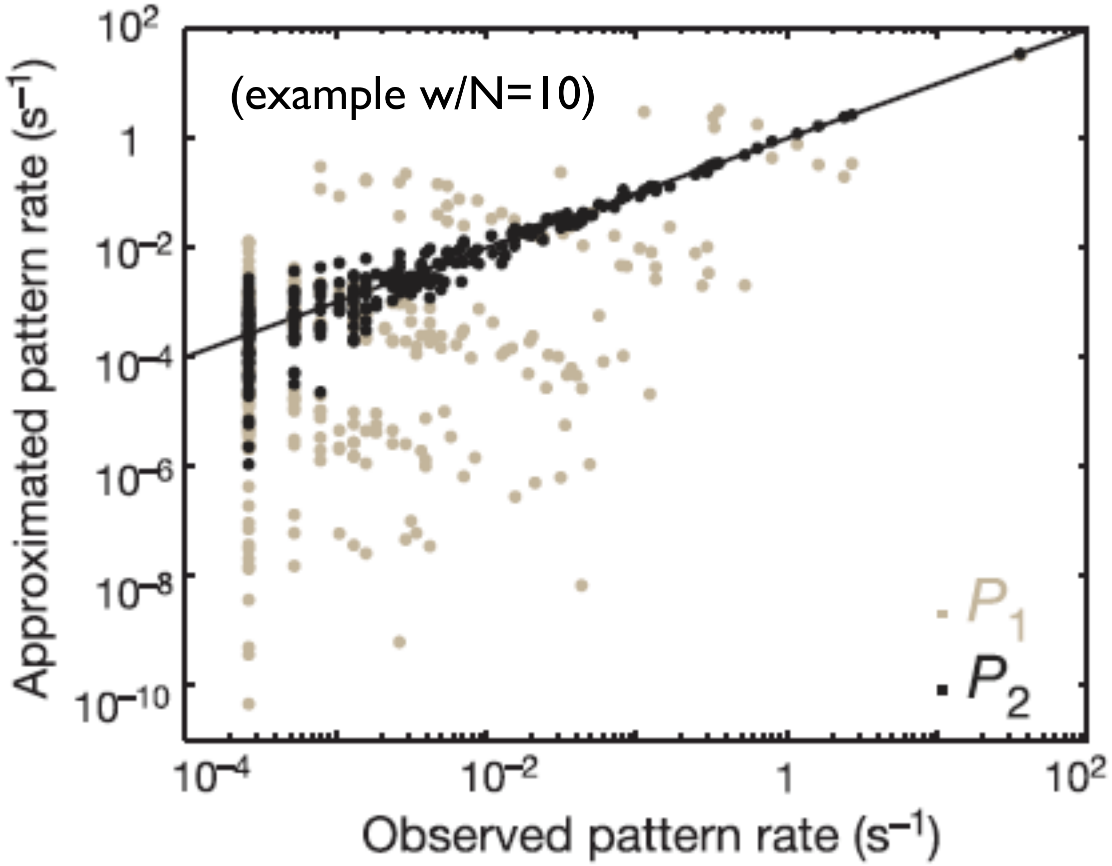



Correlations reproduced within error bars.

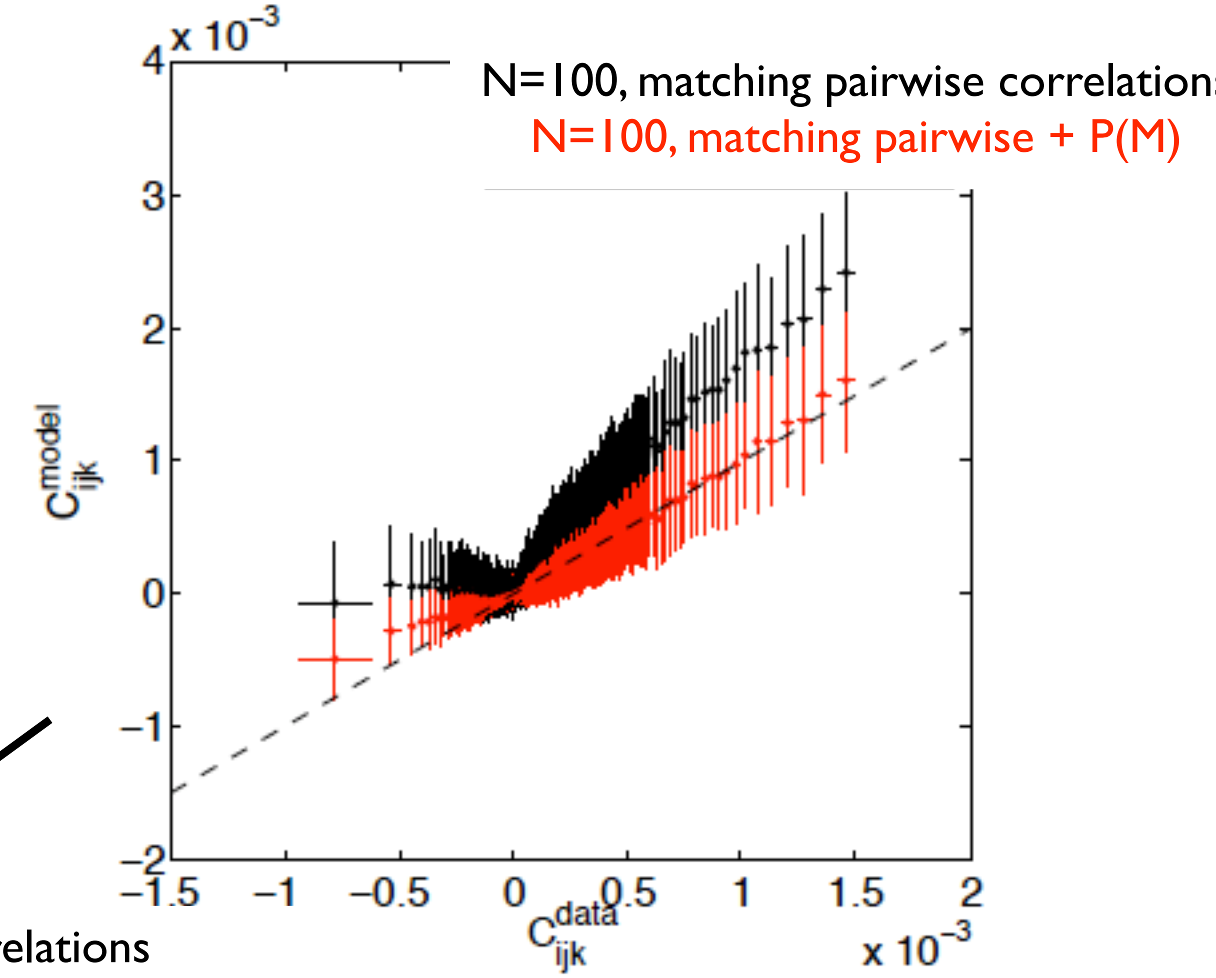


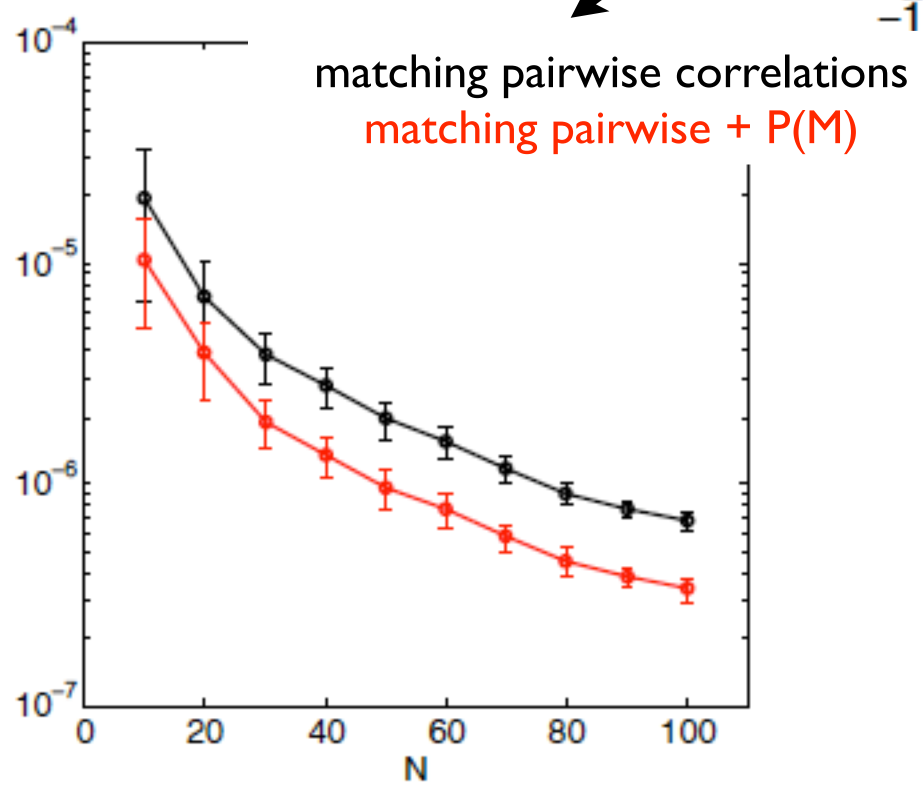



For small networks, we can test the model by checking the probability of every state.




For larger networks, we can check connected higher order (e.g. 3-cell) correlations.



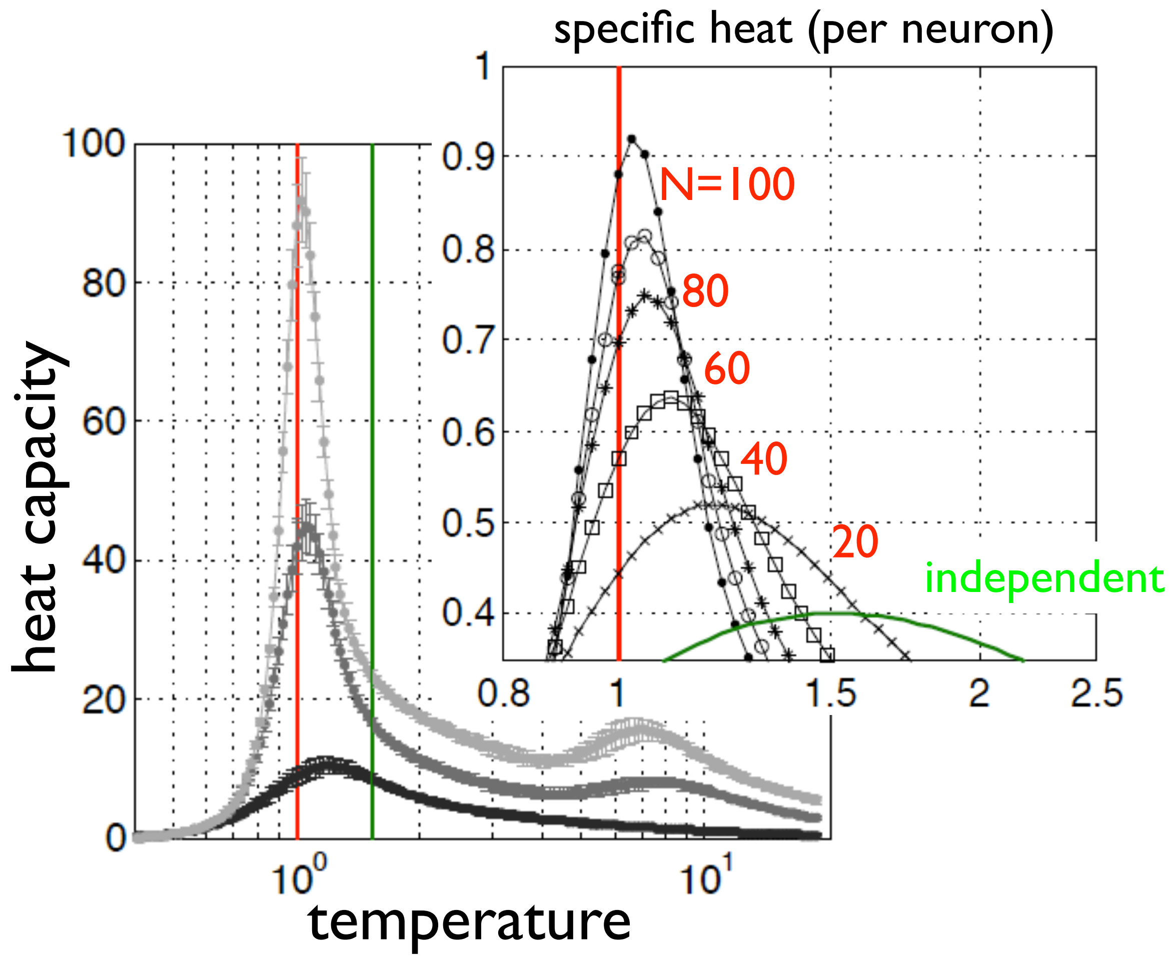




mean absolute error of predicted 3-cell correlations

Seems to be working even better for larger N.

**Where are we in parameter space?**

One direction in parameter space corresponds to changing temperature ... let's try this one:




The system is poised very close to a point in parameter space where the specific heat is maximized - a critical point.

Can we do experiments to show that the system adapts to hold itself at criticality?

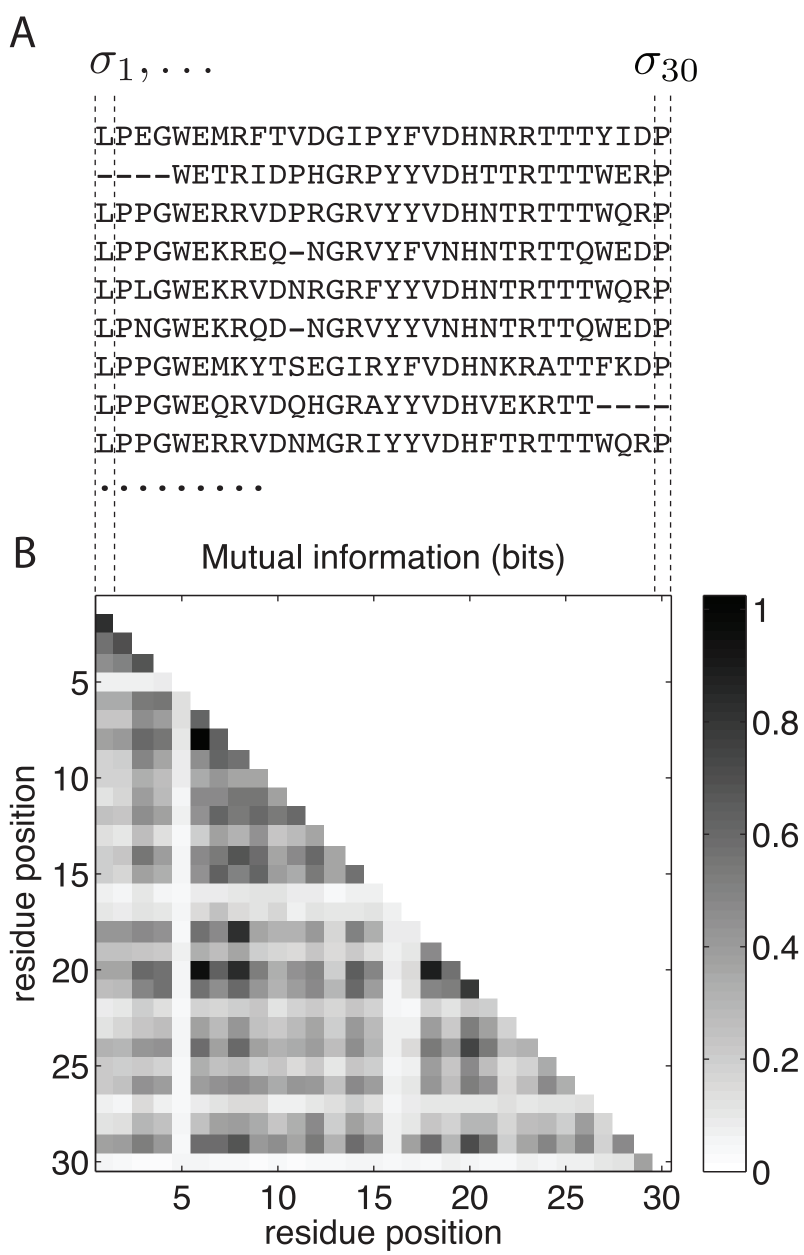
specific heat = variance of energy = variance of log(probability)

Having this be large is exactly the opposite of the usual criteria for efficient coding (!).

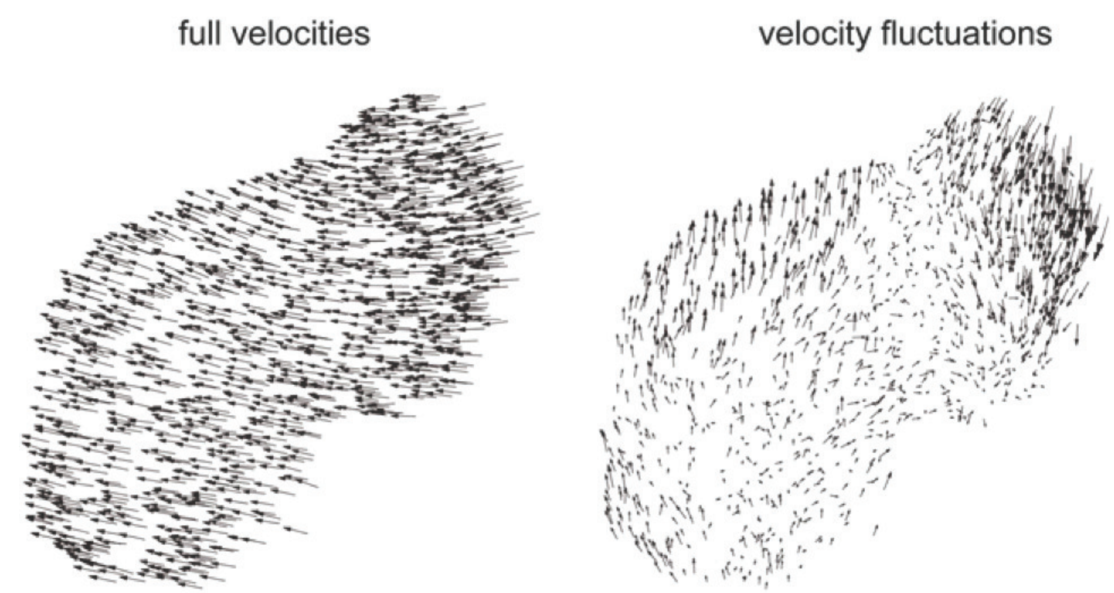
Instead, does operation near the critical point maximize the dynamic range for representing surprise?

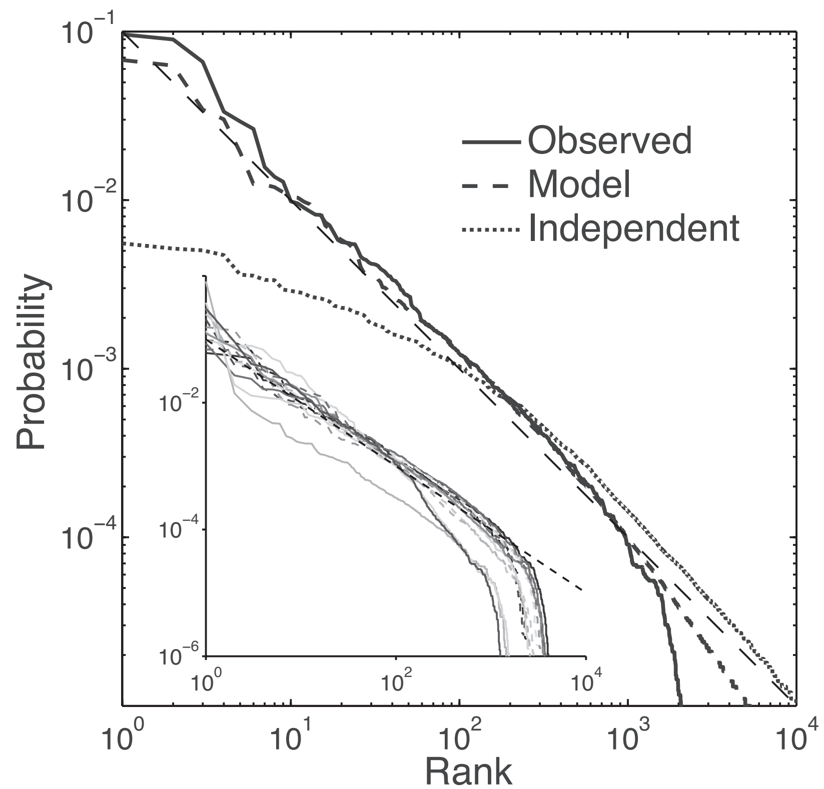
# Can use the same strategy to make models for ...

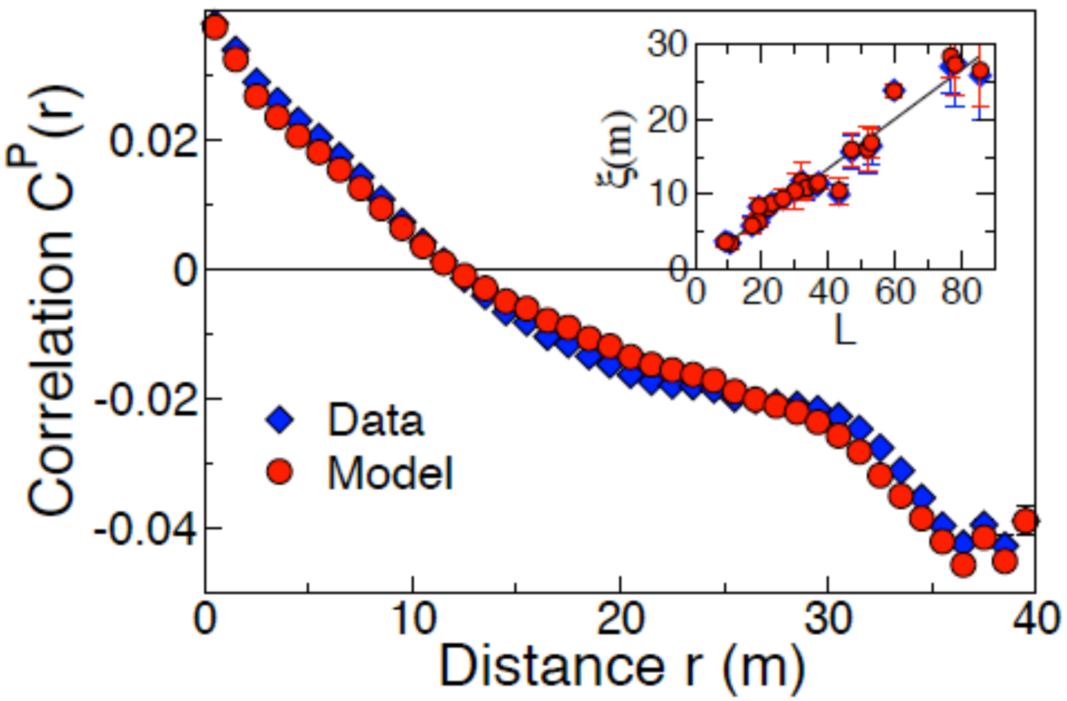
the distribution of amino acids in a family of proteins,




or the flight velocities of birds in a flock.






For protein families (here, antibodies), look at the "Zipf plot." This is S vs. E, turned on its side; unit slope implies S = E (again!).

For birds, look at the correlations directly in real space (as usual in stat mech).

ALL of these, as with the specific heat in our neural network, are signatures that the real system is operating near a critical point in its parameter space.