Statistical mechanics for a network of real neurons

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

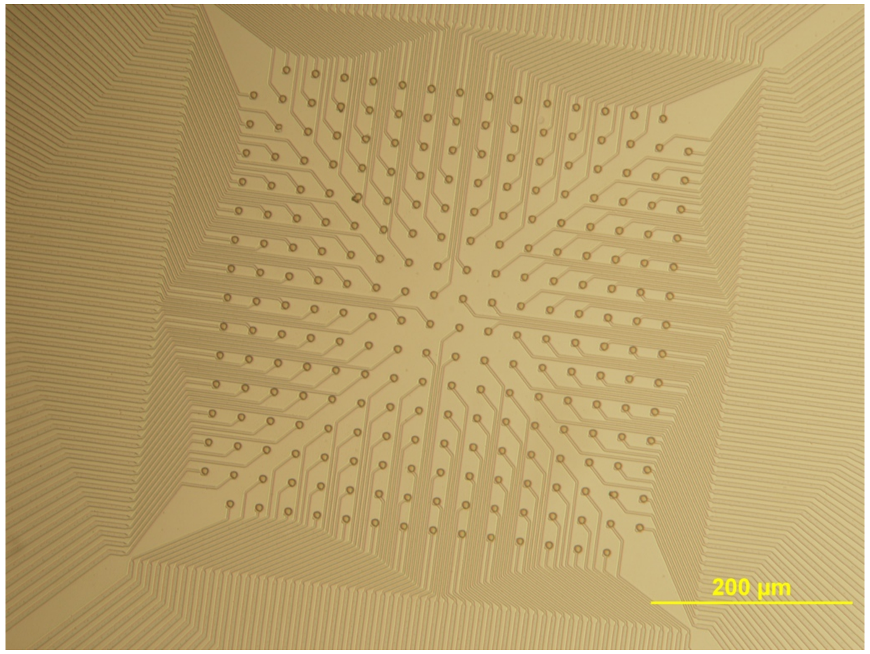
<u>Theoretical collaborators</u>:
T Mora (CNRS/ENS Paris)
E Schneidman (Weizmann)
S Still (U. Hawaii)
G Tkacik (IST Austria)

<u>Experimental collaborators</u>:
D Amodei
MJ Berry II
O Marre
R Segev (Ben-Gurion U.)

Network information and connected correlations.
E Schneidman, S Still, MJ Berry II & W Bialek, *Phys Rev Lett* <u>91</u>, 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* <u>440</u>, 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).
Spin glass models for networks of real neurons.
G Tkacik, E Schneidman, MJ Berry II & W Bialek, arXiv:0912.5409 [q-bio.NC] (2009).
Are biological systems poised at criticality?
T Mora & W Bialek, *J Stat Phys* in press (2011); arXiv:1012.2242 [q-bio.QM] (2010).

and, most importantly, work in progress.

With modern electrode arrays one can record the activity of > 100 neurons at once.



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

Importantly, these are almost all the ganglion cells within the radius of connectivity.

<u>State of one element of a network</u>:

spike/silence from one neuron

expression level of one gene

choice of amino acid at one site in a protein

flight direction and speed of one bird in a flock

$$\sigma_{\rm n}$$

<u>State of the entire network</u>

$$\{\sigma_{\rm n}\}$$

What do we want to know?

How do we get from one state to another?

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

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, ... 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 (!).

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

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

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

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

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

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

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

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:

$\{\langle f_\mu\rangle\} =$ firing rates $\Rightarrow E = \sum_{\rm n} h_{\rm n}\sigma_{\rm n}$

firing rates + 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 of M cells firing 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 firing together

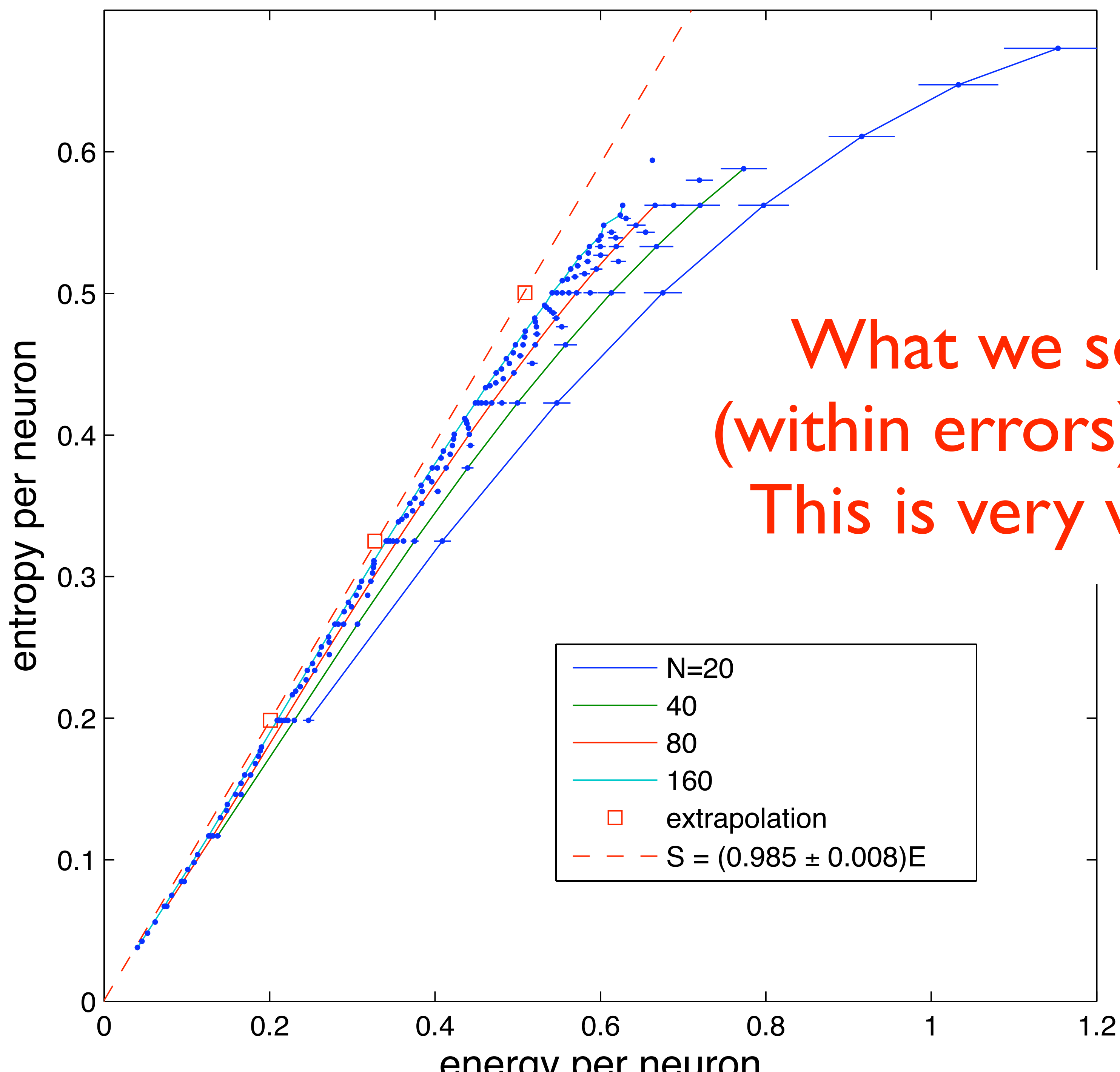
$$\Rightarrow E = V\left(\sum_{n} \sigma_{n}\right)$$

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.



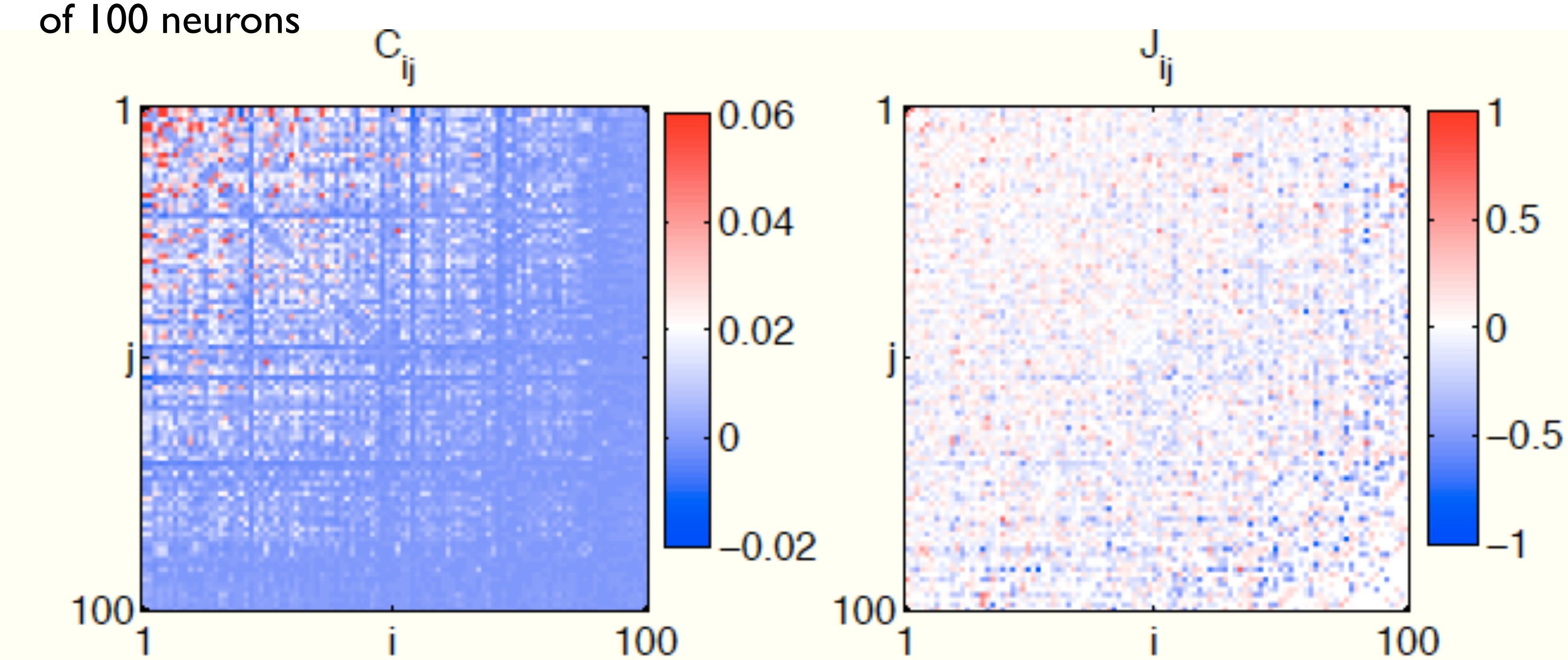
What we see is (within errors) S = E. This is very weird.

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

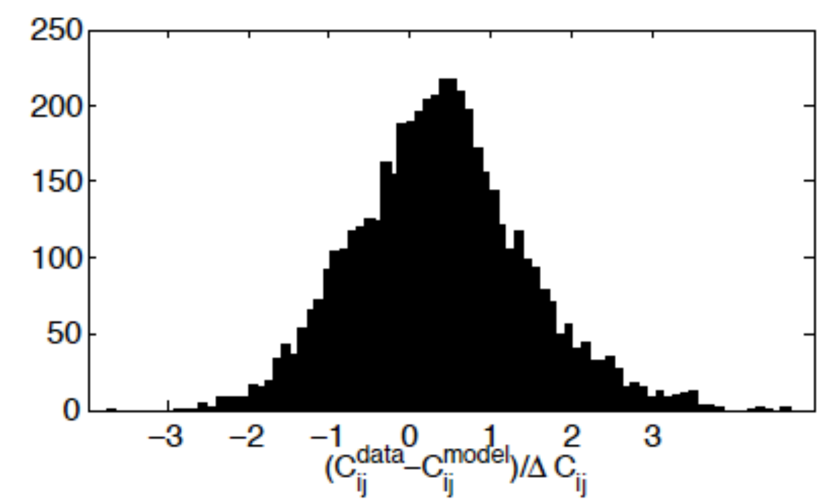
$$E = \sum_{\mathrm{n}} h_{\mathrm{n}}\sigma_{\mathrm{n}} + \frac{1}{2}\sum_{\mathrm{nm}} J_{\mathrm{nm}}\sigma_{\mathrm{n}}\sigma_{\mathrm{m}} + V\left(\sum_{\mathrm{n}}\sigma_{\mathrm{n}}\right)$$

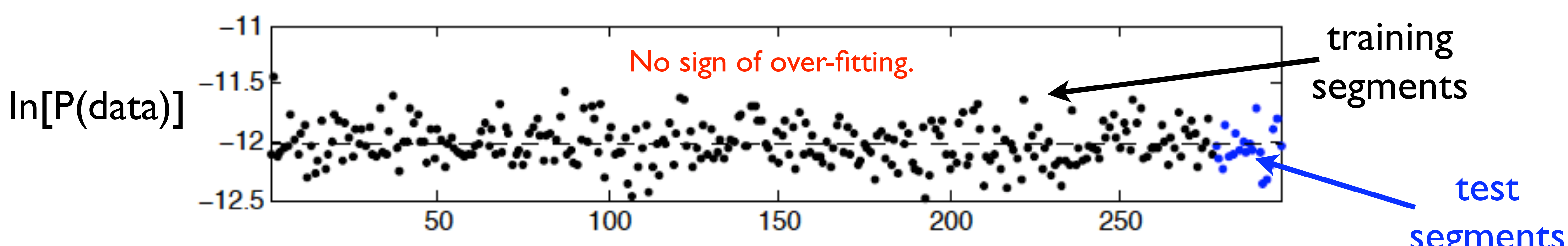
There are lots of parameters, but we can find them all from ~1 hour of data. This is a hard "inverse statistical mechanics" problem.

one example of 100 neurons

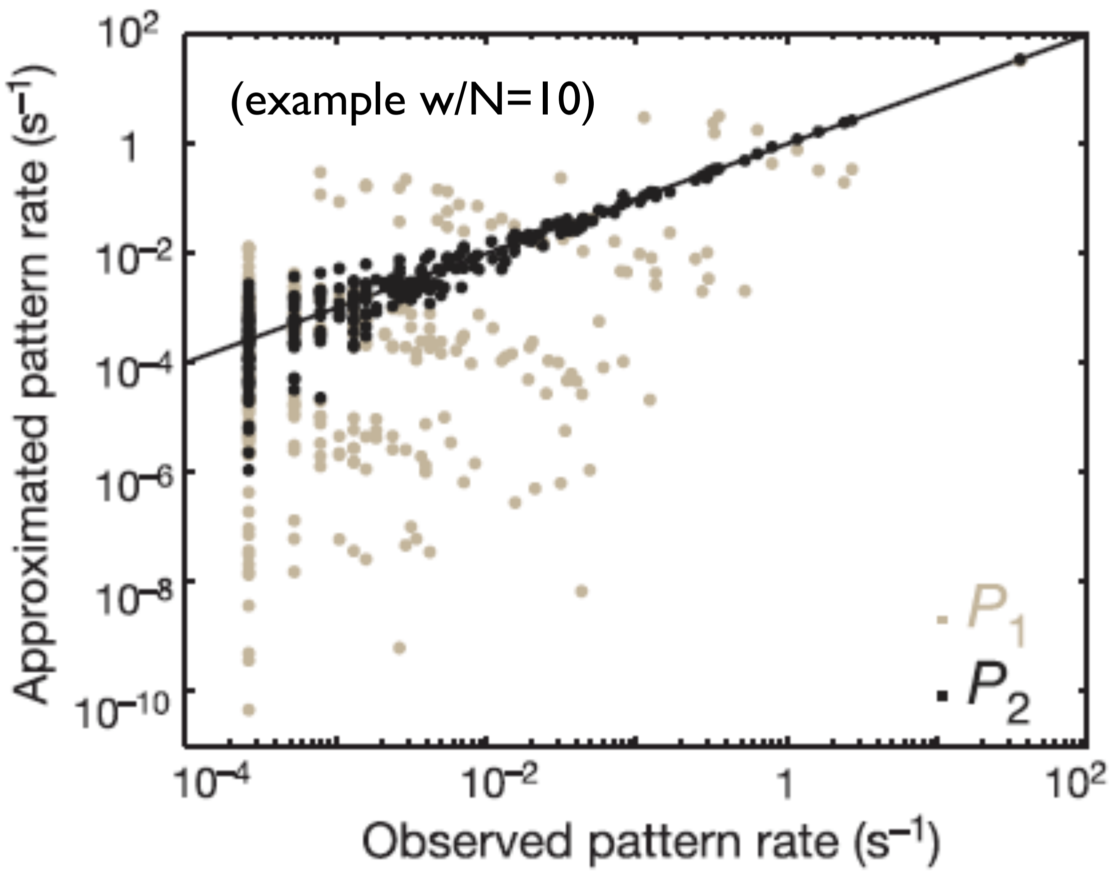


Correlations reproduced within error bars.

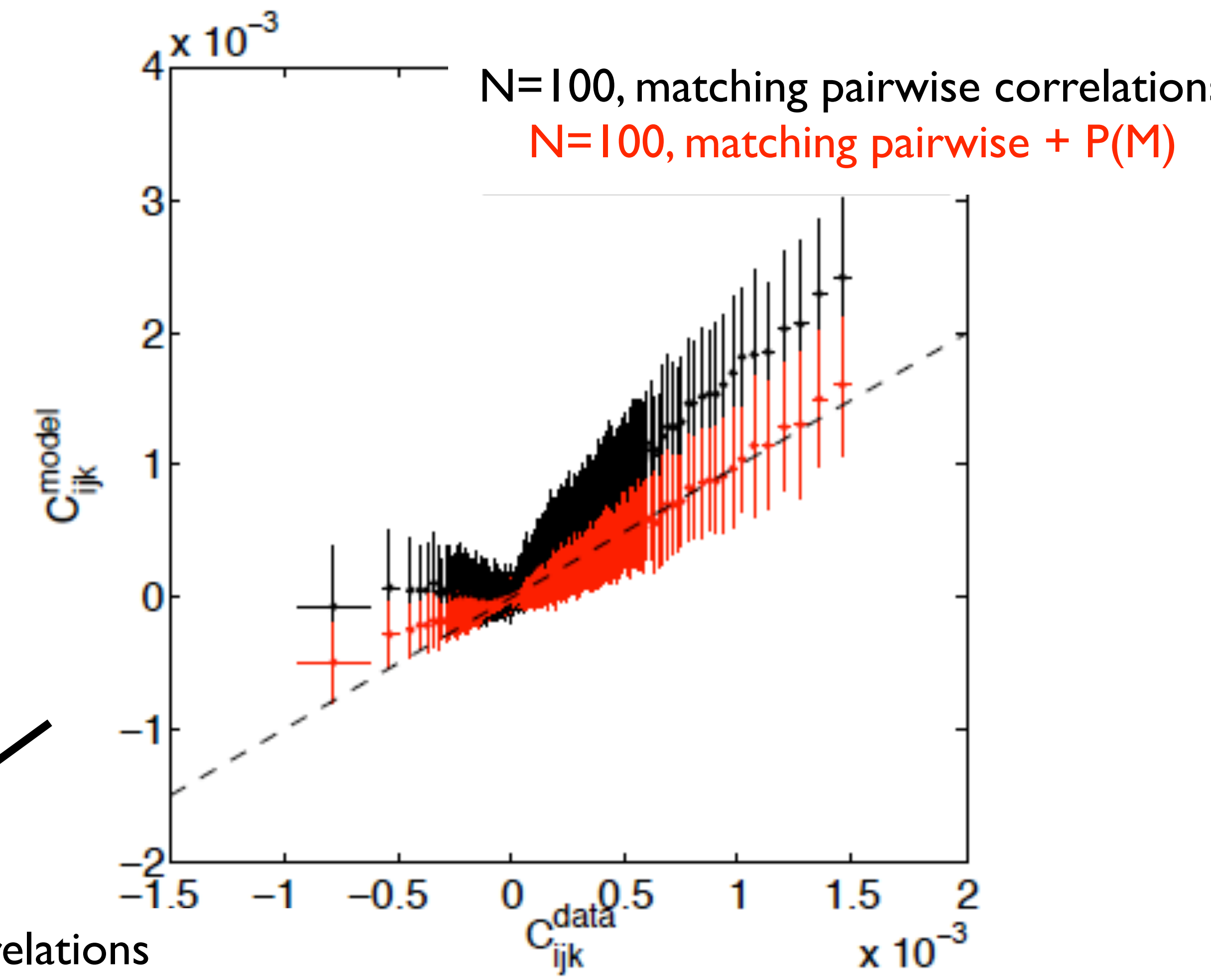




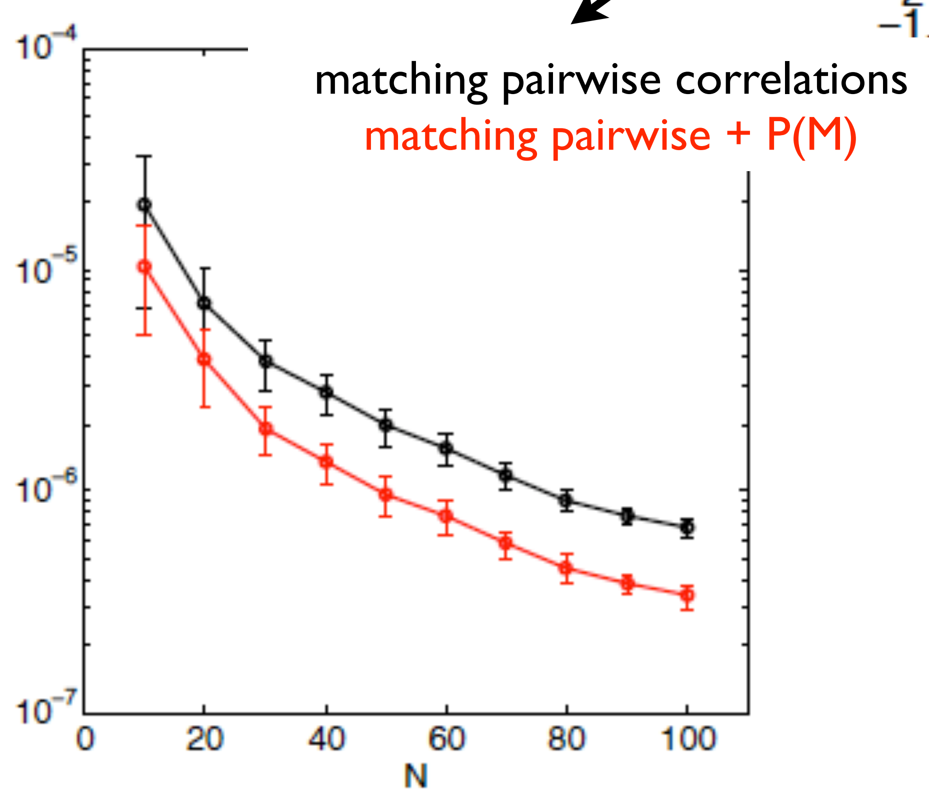
For small networks, we can test the model by checking the probability of every state.



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



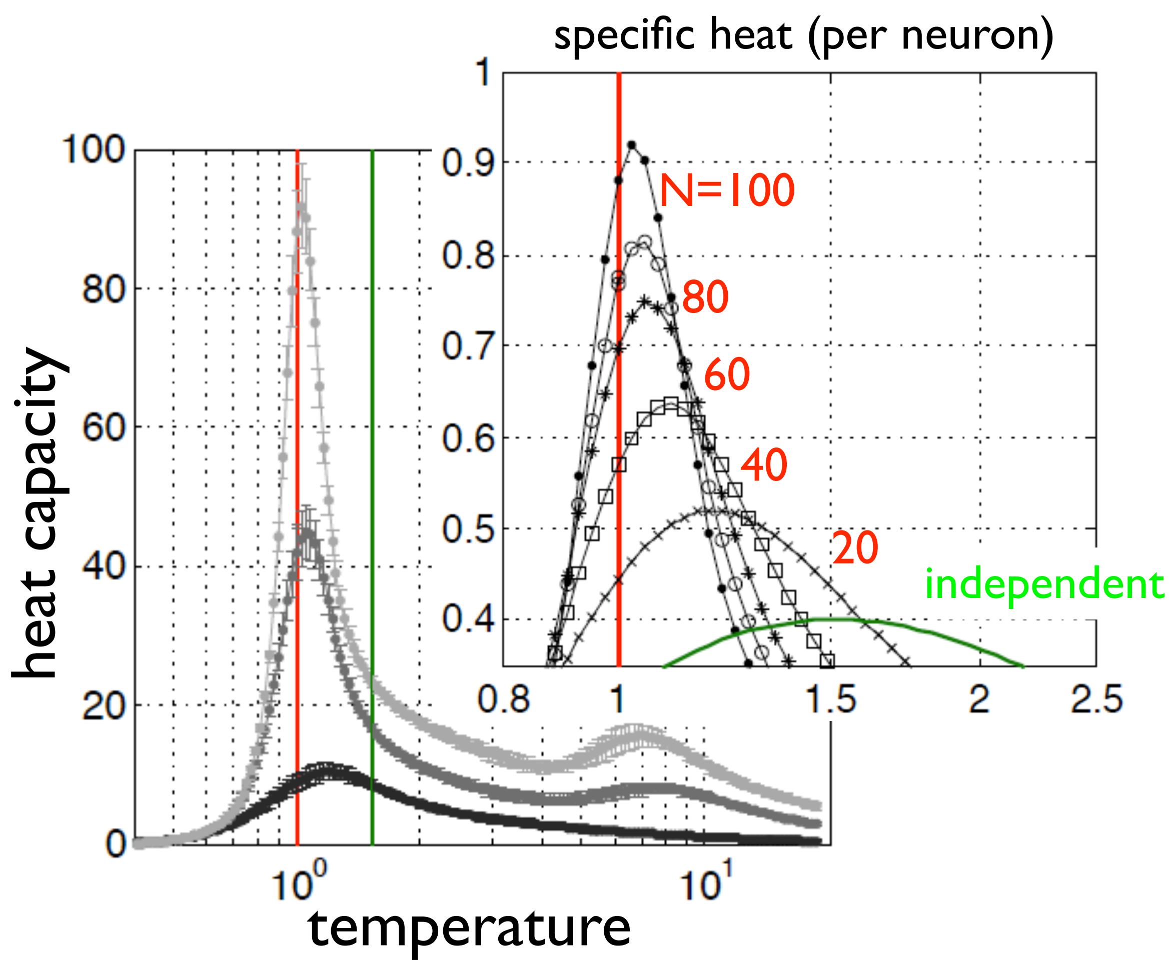
mean absolute error of predicted 3-cell correlations



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?

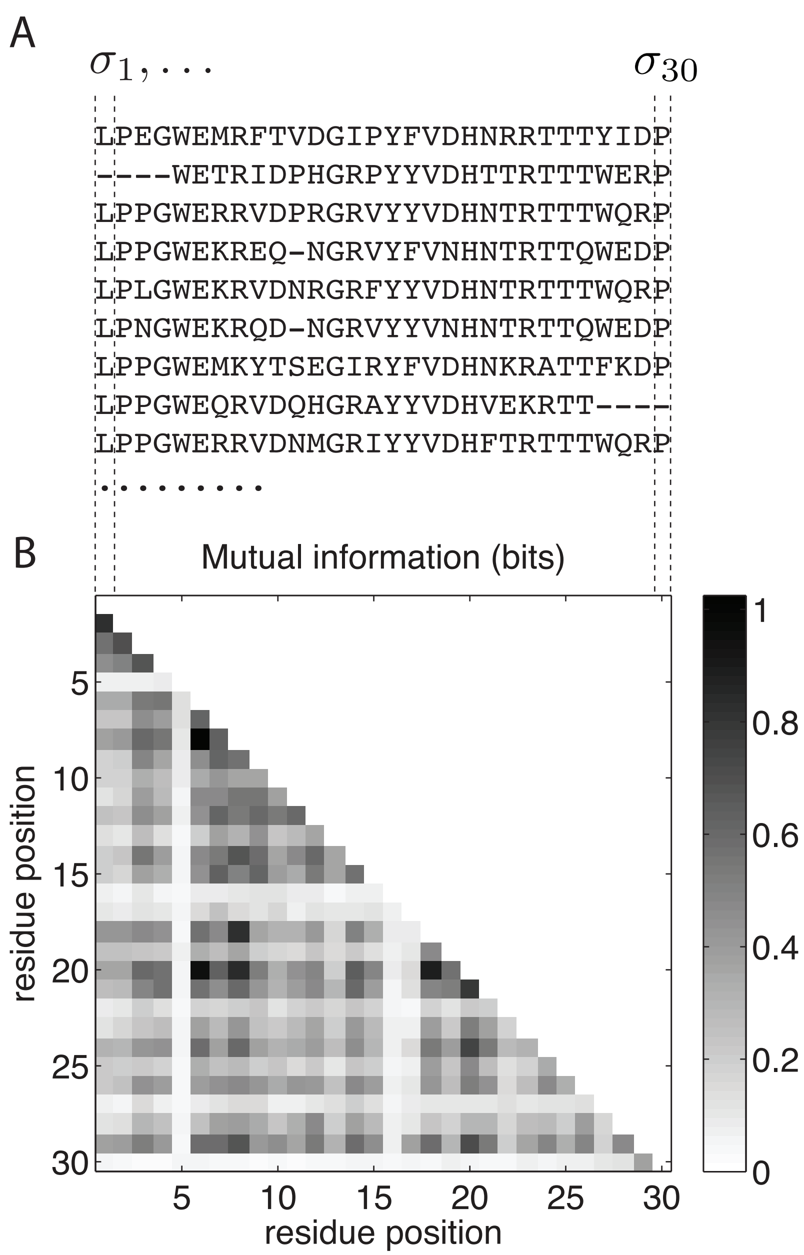
specific heat = variance of energy = variance of log(probability)

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

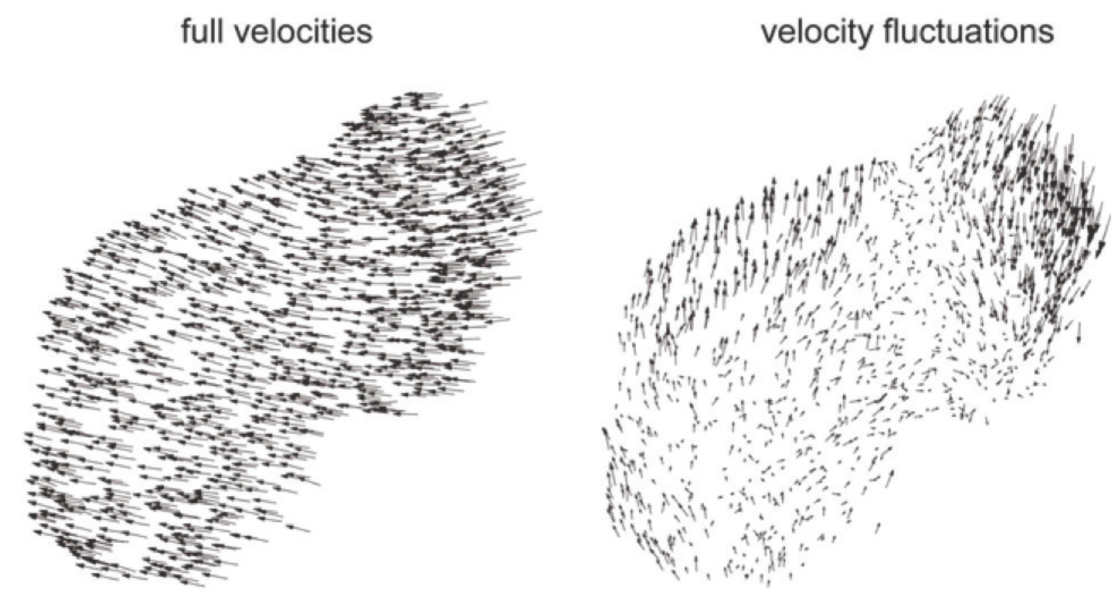
Instead, does operation near the critical point maximize the dynamic range for representing surprise?

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

the distribution of amino acids in a family of proteins,

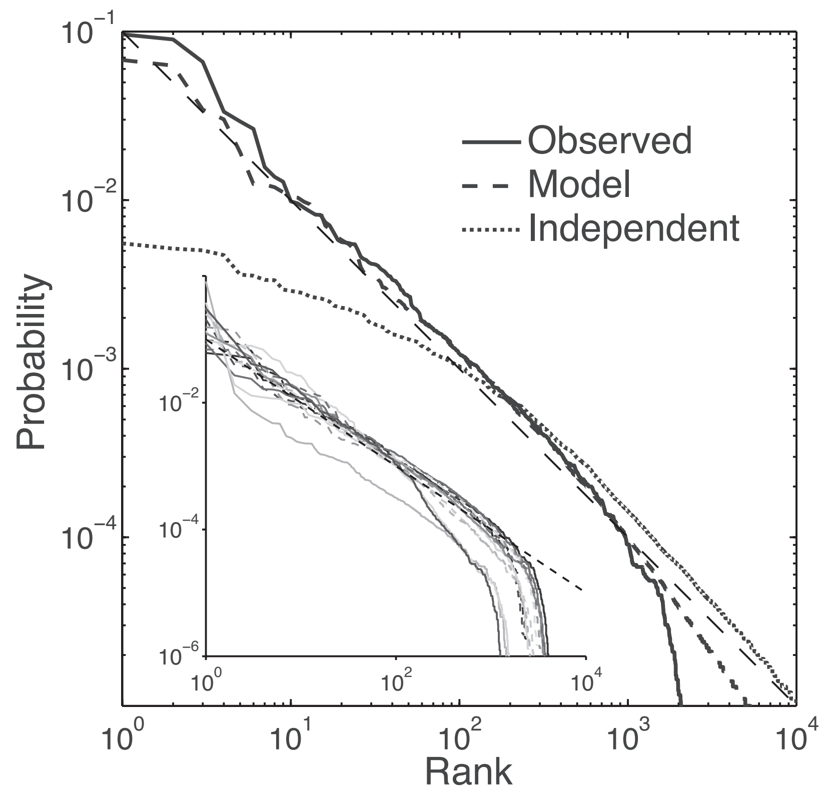


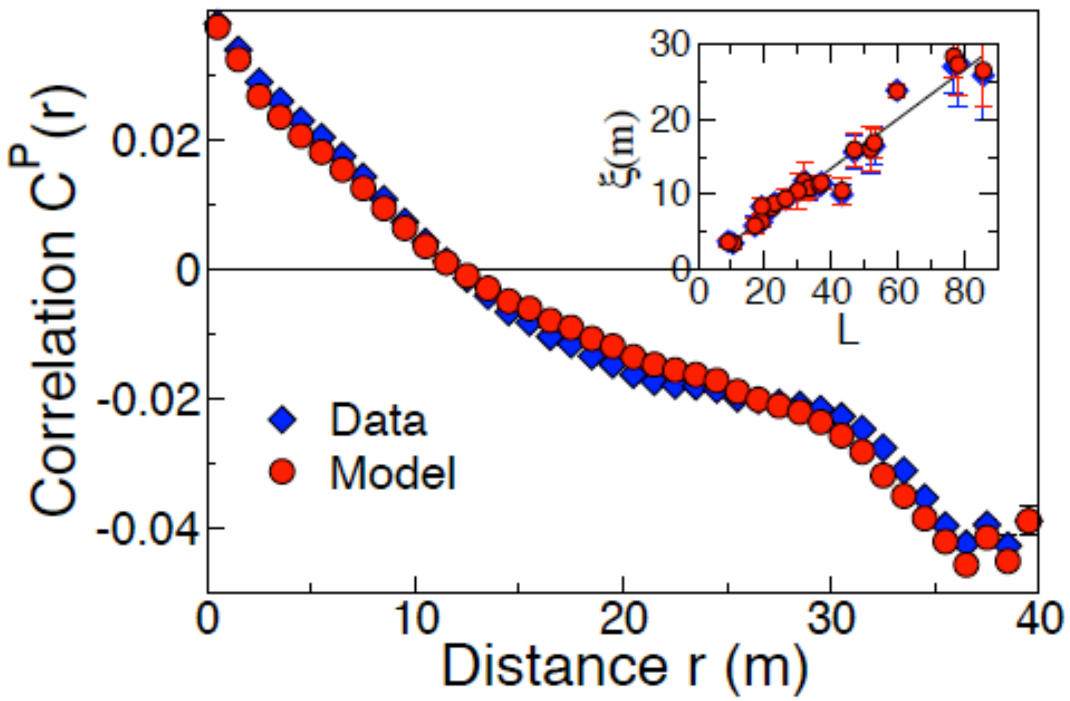
or the flight velocities of birds in a flock.



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

Statistical mechanics for a natural flock of birds.
W Bialek, A Cavagna, I Giardina, T Mora, E Silvestri, M Viale & A Walczak, arXiv.org: 1107.0604 [physics.bio-ph] (2010).





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.