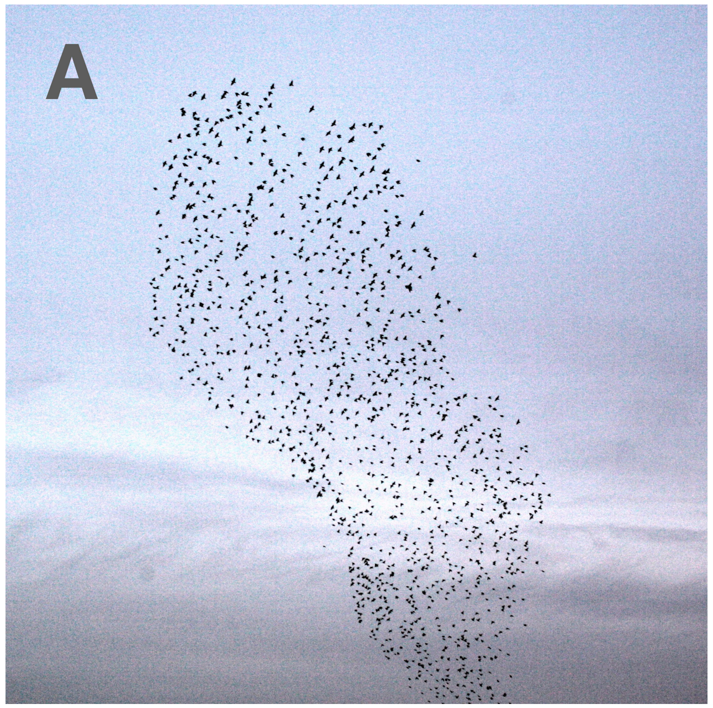
Learning the dynamics of biological networks

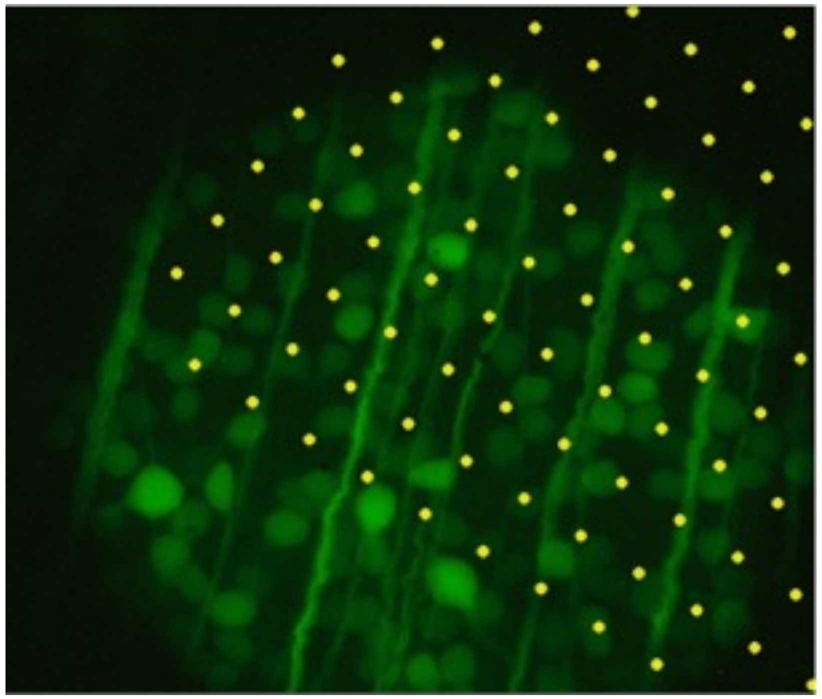
Thierry Mora
Laboratoire de physique statistique
École normale supérieure, Paris
& CNRS

ENS Paris
A. Walczak

Università Sapienza Rome
A. Cavagna
I. Giardina
O. Pohl
E. Silvestri
M. Viale

Aberdeen University
F. Ginelli





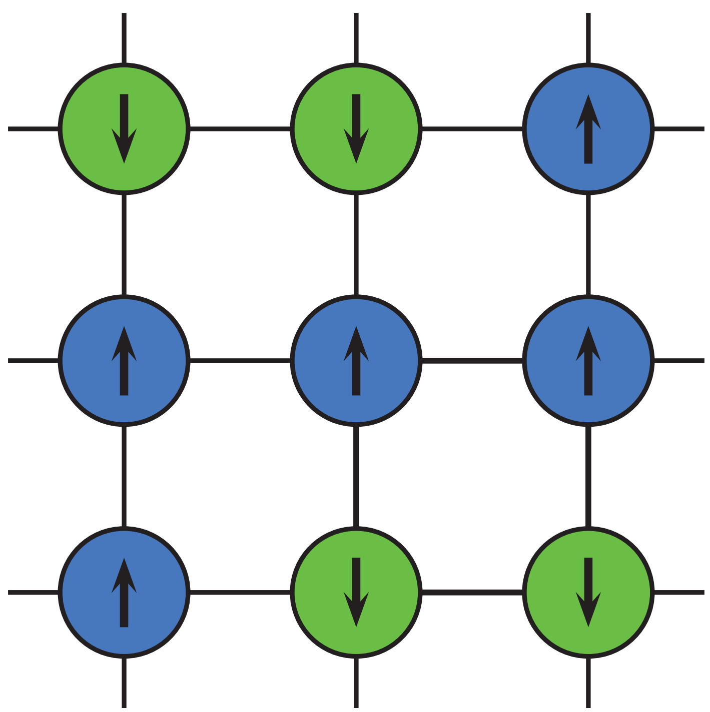
IST Austria
G. Tkacik

Vision Institute
S. Deny
O. Marre

Princeton University
W. Bialek
M. Berry

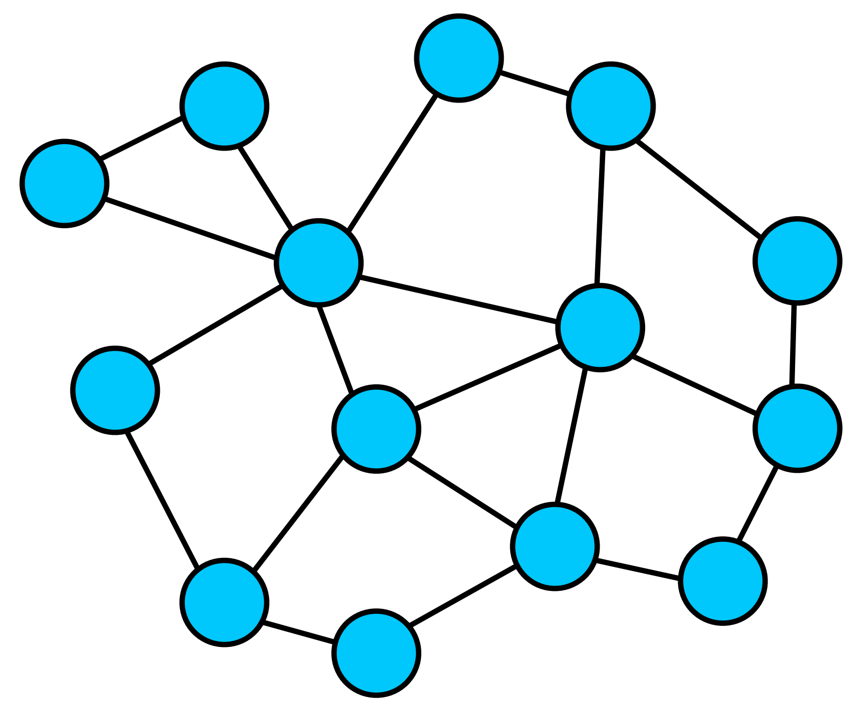
statistical mechanics as a tool to describe correlated systems

interacting spins



spontaneous magnetization

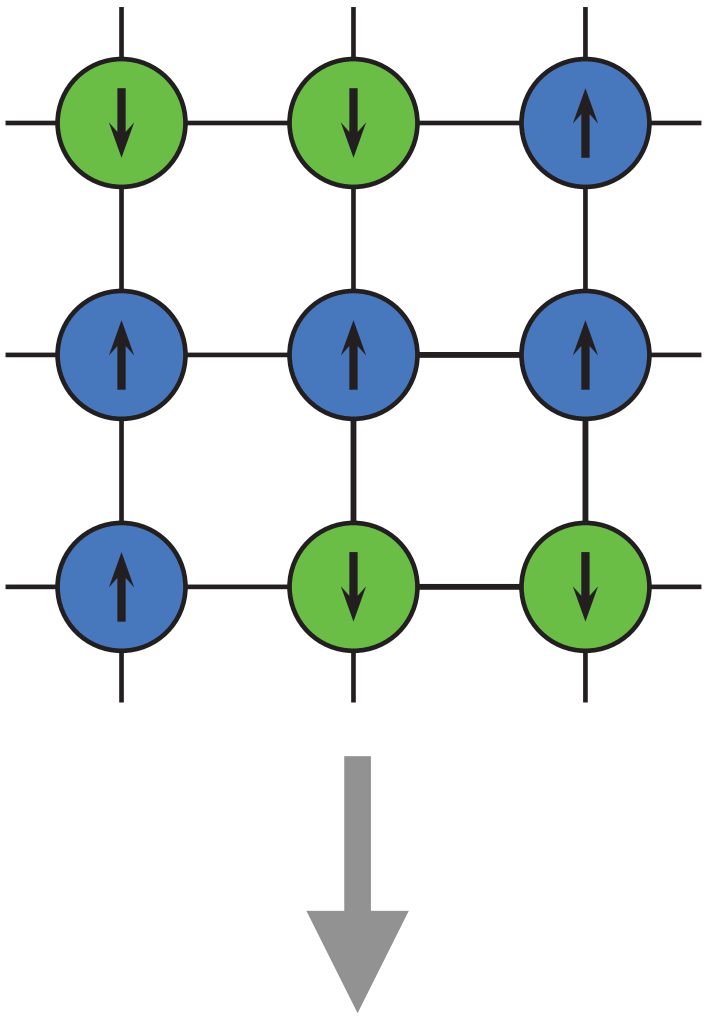
(any) interacting agents



collective behaviour

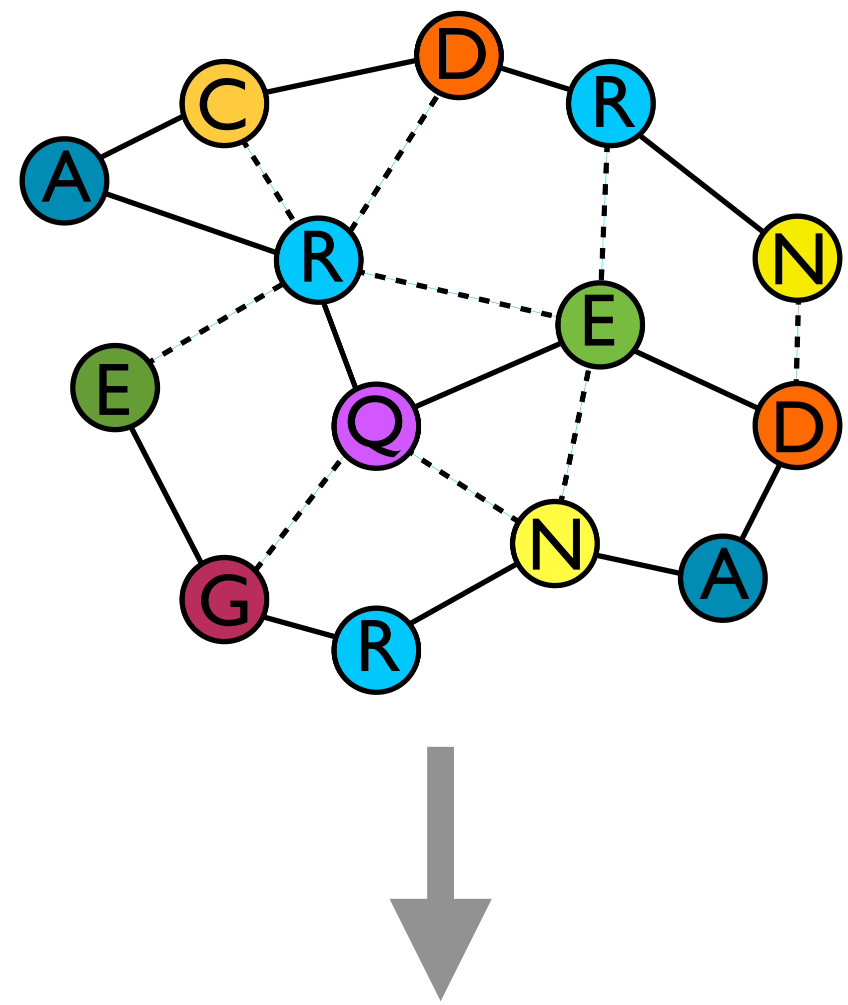
statistical mechanics as a tool to describe correlated systems

interacting spins



spontaneous magnetization

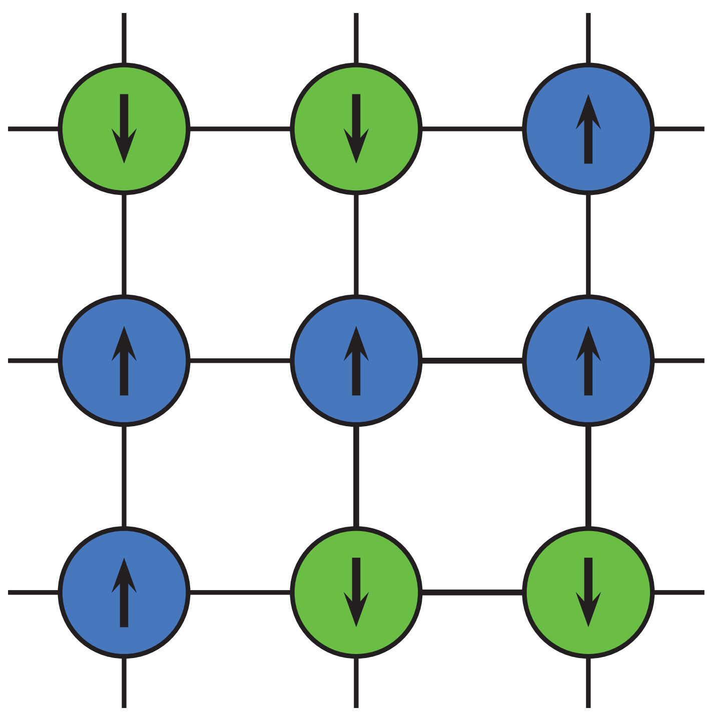
(any) interacting agents



collective behaviour

statistical mechanics as a tool to describe correlated systems

interacting spins



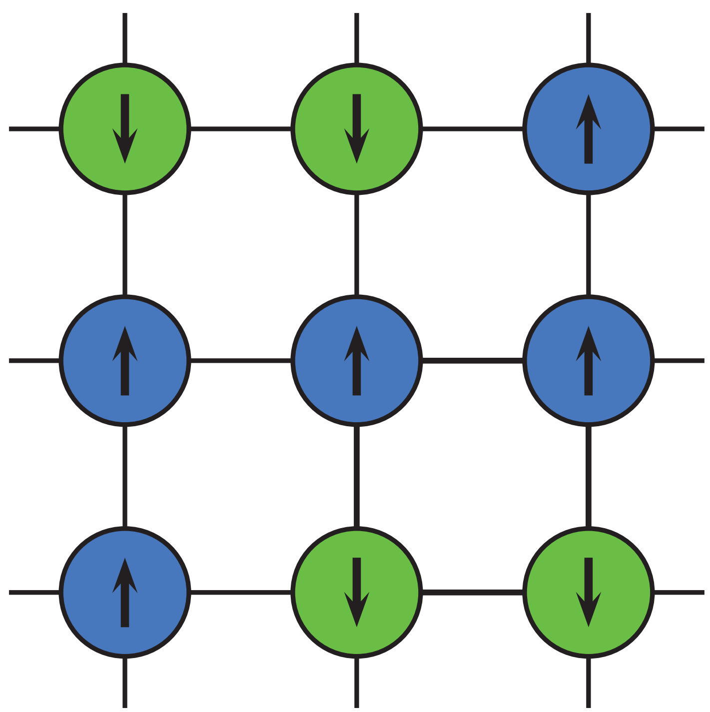
spontaneous magnetization

(any) interacting agents

collective behaviour

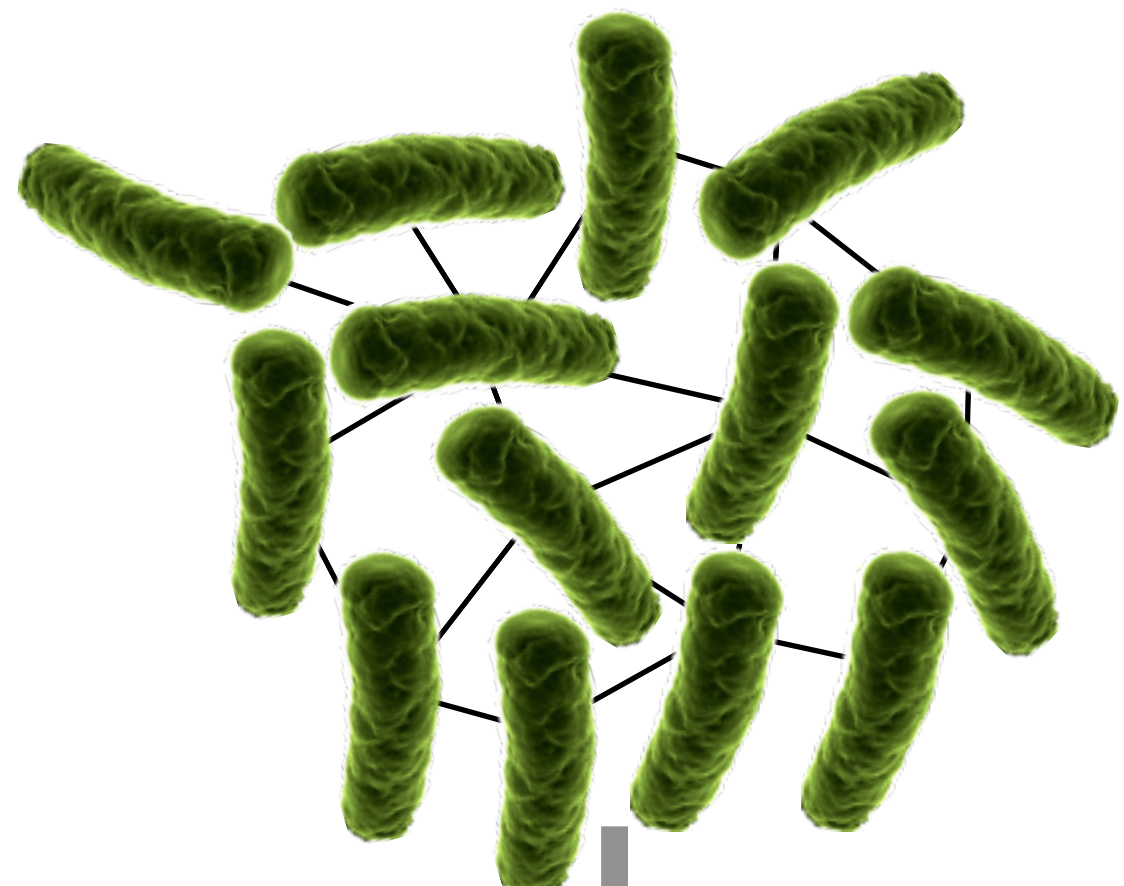
statistical mechanics as a tool to describe correlated systems

interacting spins



spontaneous magnetization

(any) interacting agents

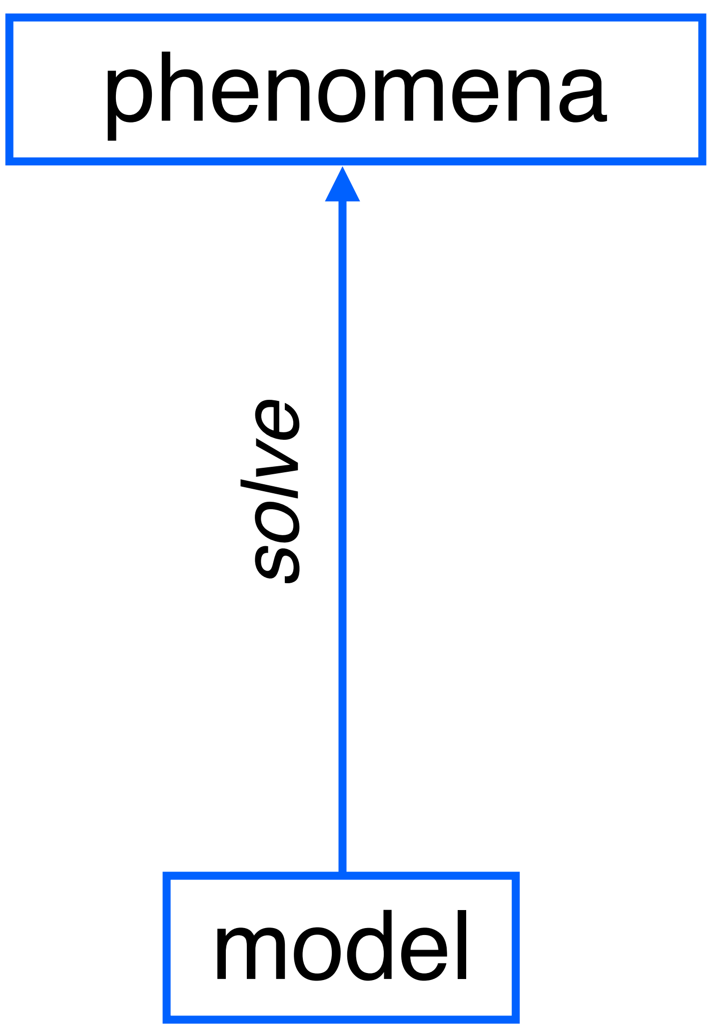


collective behaviour

two modeling approaches

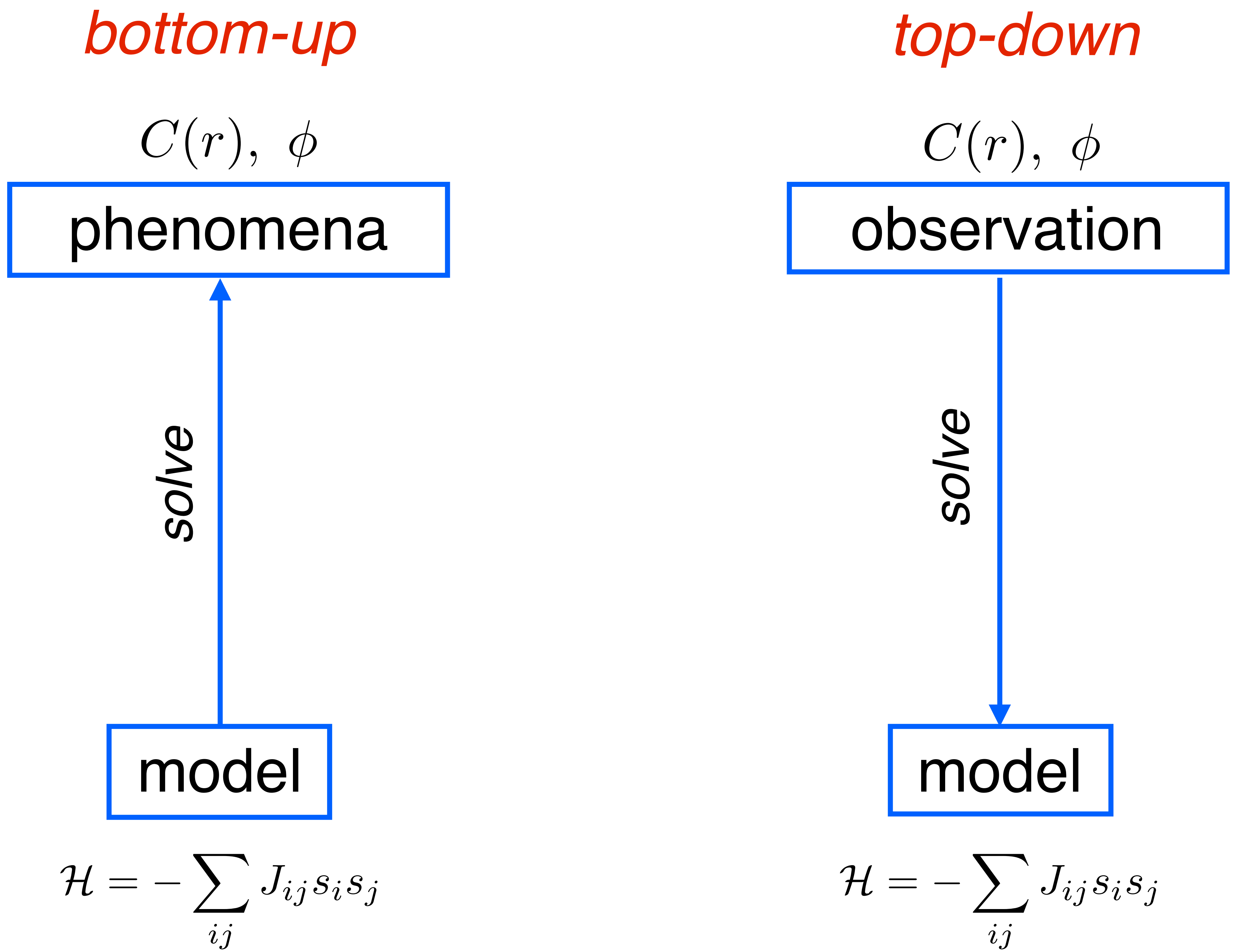
bottom-up

$C(r),\ \phi$

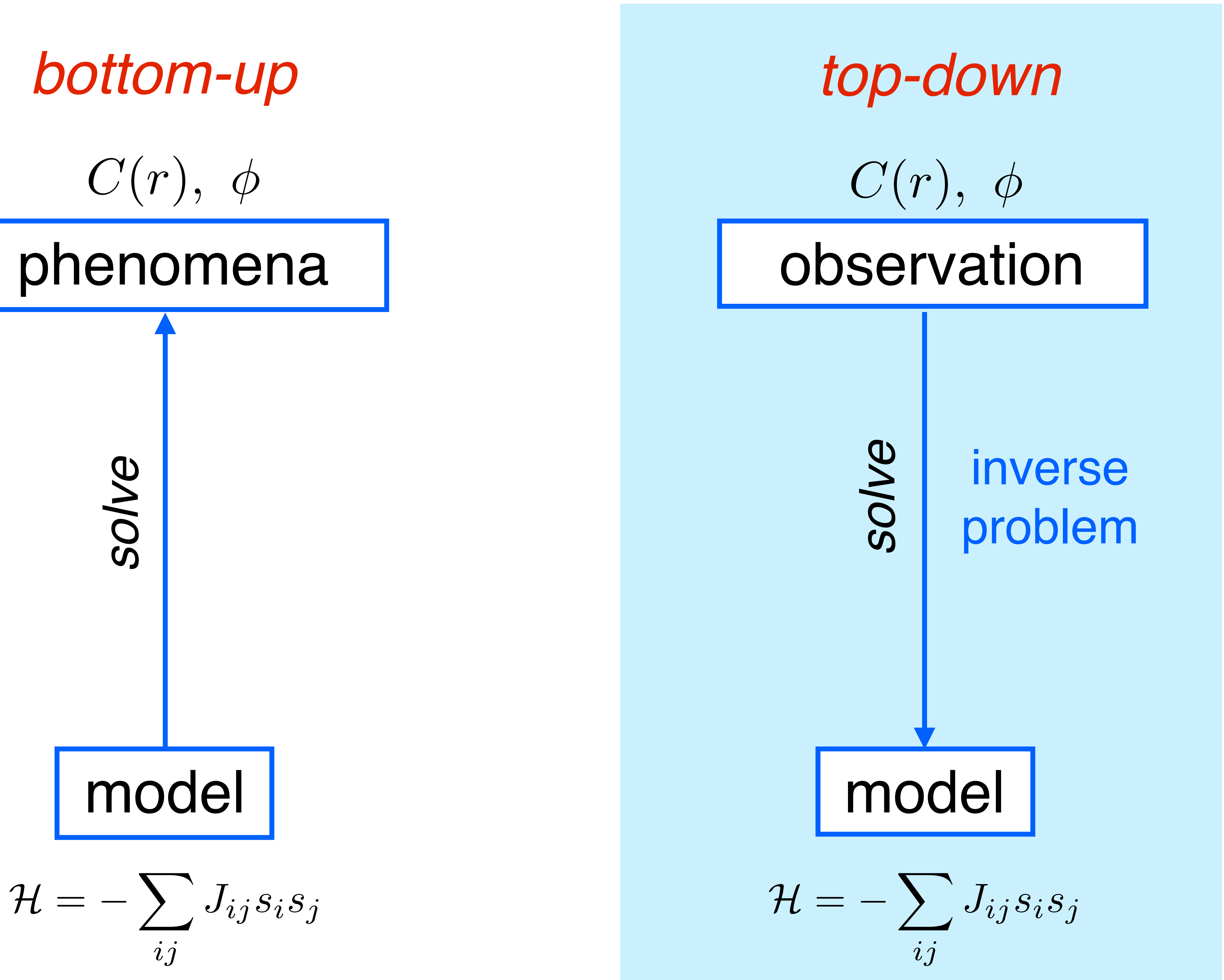


$$\mathcal{H} = -\sum_{ij} J_{ij} s_i s_j$$

two modeling approaches

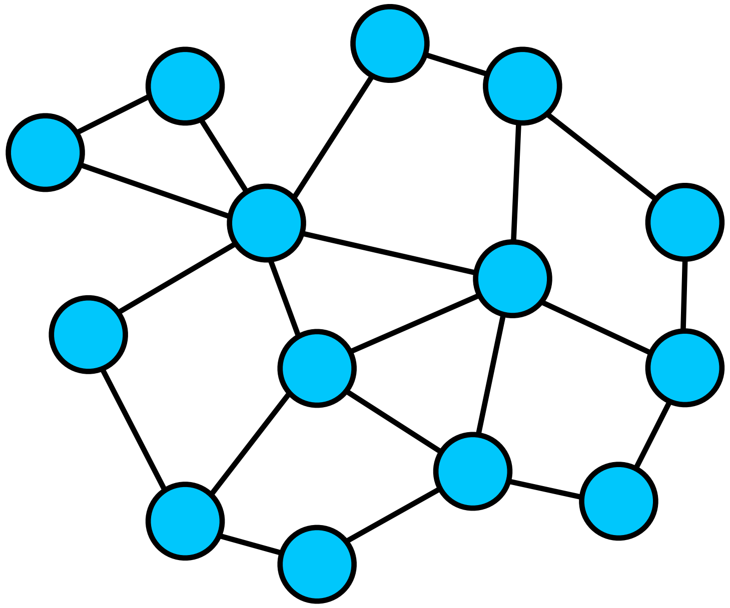


two modeling approaches



how to fit models to data: the maximum entropy approach

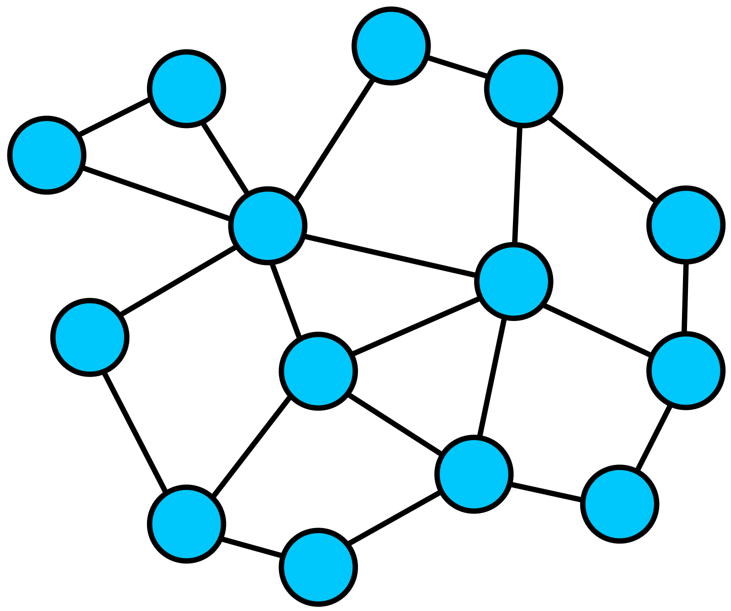
$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$



how to fit models to data: the maximum entropy approach

- N agents / units described by a variable σ

$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$

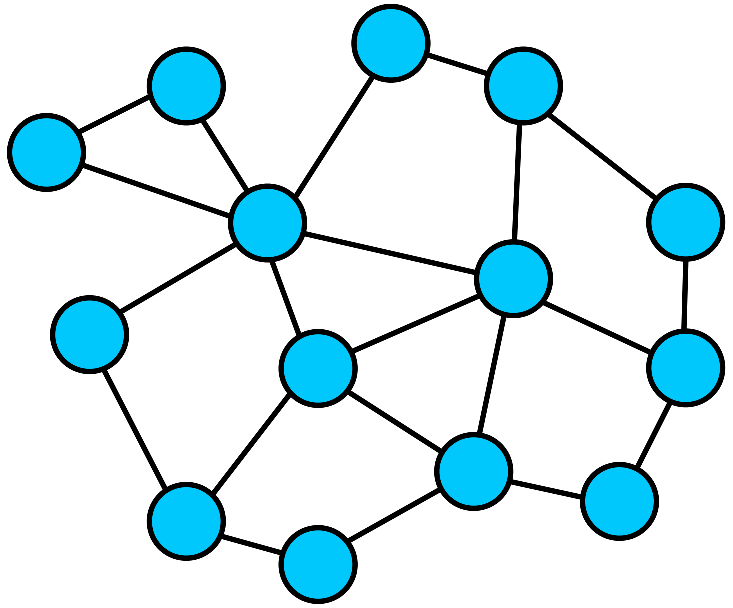


how to fit models to data: the maximum entropy approach

- N agents / units described by a variable σ

$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$

- Maximize the entropy $S = -\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \log P(\boldsymbol{\sigma})$

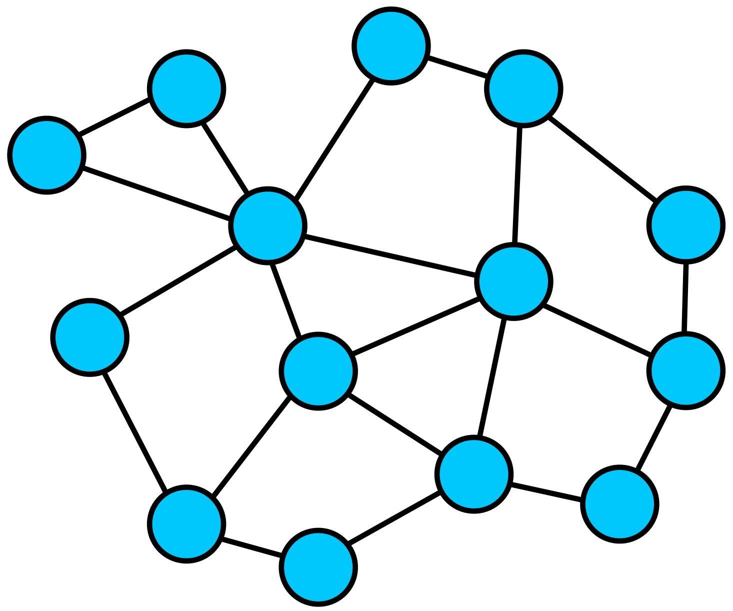


how to fit models to data: the maximum entropy approach

- N agents / units described by a variable σ

$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$

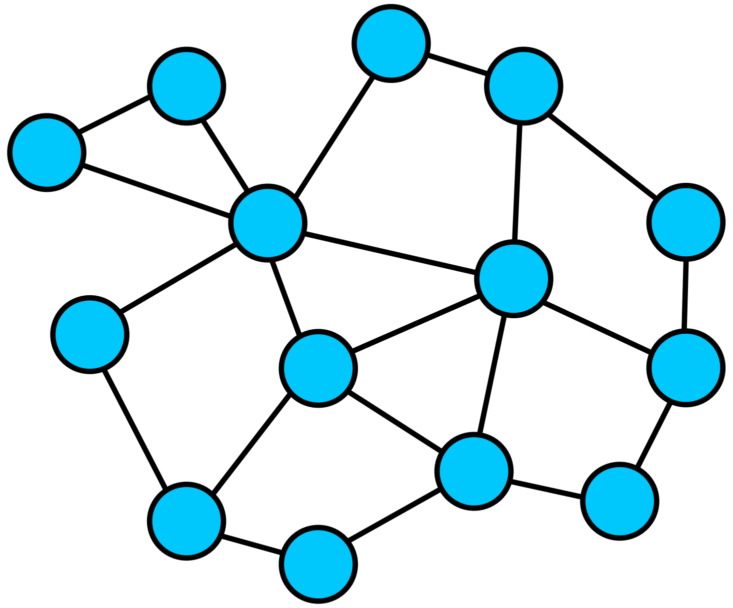
- Maximize the entropy $S = -\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \log P(\boldsymbol{\sigma})$
- under the constraint that observables $\mathcal{O}_1, \mathcal{O}_2, \ldots$ have the same average as the data



$$\langle \mathcal{O}_a \rangle_{\rm model} = \langle \mathcal{O}_a \rangle_{\rm data}$$

$\langle \mathcal{O}_a \rangle$ is typically a moment, e.g. $\langle \sigma_i \rangle, \langle \sigma_i \sigma_j \rangle$

how to fit models to data: the maximum entropy approach



- N agents / units described by a variable σ

$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$

- Maximize the entropy $S = -\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \log P(\boldsymbol{\sigma})$
- under the constraint that observables $\mathcal{O}_1, \mathcal{O}_2, \ldots$ have the same average as the data

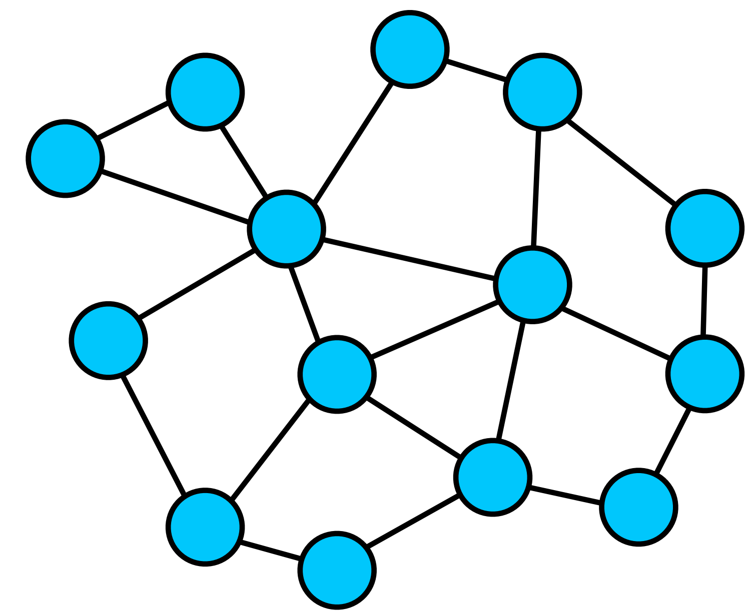
$$\langle \mathcal{O}_a \rangle_{\text{model}} = \langle \mathcal{O}_a \rangle_{\text{data}}$$

$\langle \mathcal{O}_a \rangle$ is typically a moment, e.g. $\langle \sigma_i \rangle, \langle \sigma_i \sigma_j \rangle$

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp \Bigg[\underbrace{\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})}_{-E/k_B T} \Bigg]$$

e.g. $$P(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_i h_i \sigma_i + \sum_{ij} \sigma_i \sigma_j}$$

how to fit models to data: the maximum entropy approach



- N agents / units described by a variable σ

$$\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$

- Maximize the entropy $S = -\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) \log P(\boldsymbol{\sigma})$

- under the constraint that observables $\mathcal{O}_1, \mathcal{O}_2, \ldots$ have the same average as the data

$$\langle \mathcal{O}_a \rangle_{\mathrm{model}} = \langle \mathcal{O}_a \rangle_{\mathrm{data}}$$

$\langle \mathcal{O}_a \rangle$ is typically a moment, e.g. $\langle \sigma_i \rangle, \langle \sigma_i \sigma_j \rangle$

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp \Big[\underbrace{\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})}_{-E/k_B T} \Big]$$

e.g. $$P(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_i h_i \sigma_i + \sum_{ij} \sigma_i \sigma_j}$$

(disordered) Ising model!

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

- Bayes rule

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

- Bayes rule

$$P(\{J_a\}|\text{data}) = \frac{P(\text{data}|\{J_a\})P(\{J_a\})}{P(\text{data})}$$

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

- Bayes rule

$$P(\{J_a\}|\text{data}) = \frac{P(\text{data}|\{J_a\})P(\{J_a\})}{P(\text{data})} \propto \left[\prod_{m=1}^{M} P(\boldsymbol{\sigma}^m|\{J_a\})\right] P(\{J_a\})$$

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

- Bayes rule

over datapoints

$$P(\{J_a\}|\text{data}) = \frac{P(\text{data}|\{J_a\})P(\{J_a\})}{P(\text{data})} \propto \left[\prod_{m=1}^{M} P(\boldsymbol{\sigma}^m|\{J_a\})\right] P(\{J_a\})$$

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

- Bayes rule

over datapoints

flat prior

$$P(\{J_a\}|\text{data}) = \frac{P(\text{data}|\{J_a\})P(\{J_a\})}{P(\text{data})} \propto \left[\prod_{m=1}^{M} P(\boldsymbol{\sigma}^m|\{J_a\})\right] \cancel{P(\{J_a\})}$$

maximum likelihood formulation

- Given the functional form

$$P(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left[\sum_a J_a \mathcal{O}_a(\boldsymbol{\sigma})\right]$$

what parameters J_a explain the data best?

- Bayes rule

over datapoints　　flat prior

$$P(\{J_a\}|\text{data}) = \frac{P(\text{data}|\{J_a\})P(\{J_a\})}{P(\text{data})} \propto \left[\prod_{m=1}^{M} P(\boldsymbol{\sigma}^m|\{J_a\})\right] P(\{J_a\})$$

$$\frac{\partial \log P}{\partial J_a} = 0 \;\Rightarrow -M\frac{\partial \log Z}{\partial J_a} + \sum_a \sum_{m=1}^{M} \mathcal{O}_a(\boldsymbol{\sigma}^m) = 0$$ (maximum likelihood)

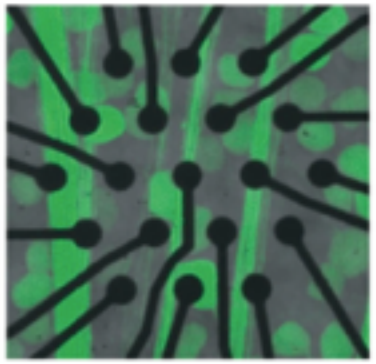
$$\Rightarrow \boxed{\langle \mathcal{O}_a(\boldsymbol{\sigma})\rangle_{\text{model}} = \langle \mathcal{O}_a(\boldsymbol{\sigma})\rangle_{\text{data}}}$$

satisfies the constaints

maximum entropy in biology

- collective activity of neural populations

  Schneidman et al. Nature 2006
  Shlens et al J. Neuroscience 2006
  Tang et al J. Neuroscience 2008
  Tkacik et al PLoS CP 2014



maximum entropy in biology

- collective activity of neural populations

  Schneidman et al. Nature 2006
  Shlens et al J. Neuroscience 2006
  Tang et al J. Neuroscience 2008
  Tkacik et al PLoS CP 2014

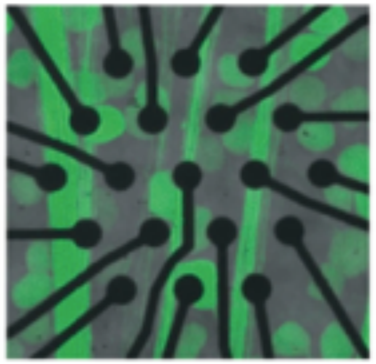
- co-variations in protein families ⟹ contact prediction

  Weigt et al. PNAS 2009; Morcos et al. PNAS 2011
  Marks et al PLoS ONE 2012; Sulkowska et al PNAS 2012

maximum entropy in biology

- collective activity of neural populations

  Schneidman et al. Nature 2006
  Shlens et al J. Neuroscience 2006
  Tang et al J. Neuroscience 2008
  Tkacik et al PLoS CP 2014



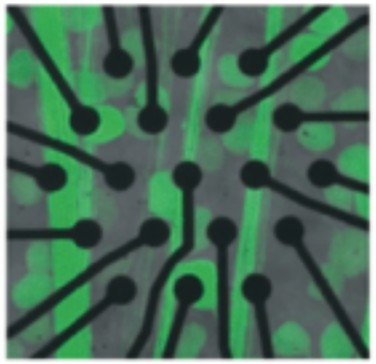
- co-variations in protein families ⟹ contact prediction

  Weigt et al. PNAS 2009; Morcos et al. PNAS 2011
  Marks et al PLoS ONE 2012; Sulkowska et al PNAS 2012

- diversity of antibody repertoires in the immune system

  Mora Walczak Callan Bialek PNAS 2010

maximum entropy in biology



- collective activity of neural populations

  Schneidman et al. Nature 2006
  Shlens et al J. Neuroscience 2006
  Tang et al J. Neuroscience 2008
  Tkacik et al PLoS CP 2014

- co-variations in protein families ⇒ contact prediction

  Weigt et al. PNAS 2009; Morcos et al. PNAS 2011
  Marks et al PLoS ONE 2012; Sulkowska et al PNAS 2012

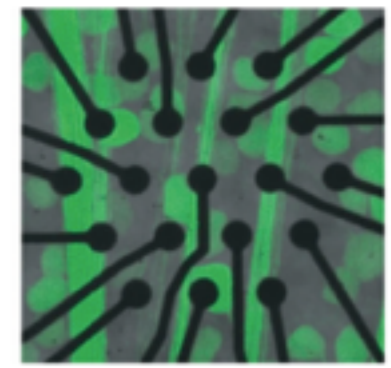
- diversity of antibody repertoires in the immune system

  Mora Walczak Callan Bialek PNAS 2010

- DNA motifs of transcription factor binding sites

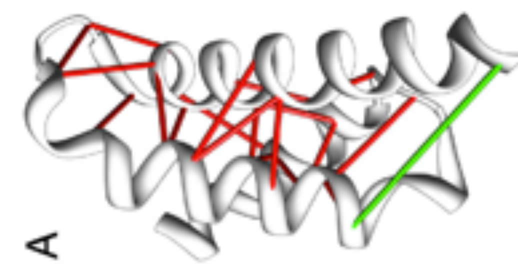
  Santolini Mora Hakim Plos ONE 2014

maximum entropy in biology



- collective activity of neural populations

  Schneidman et al. Nature 2006
  Shlens et al J. Neuroscience 2006
  Tang et al J. Neuroscience 2008
  Tkacik et al PLoS CP 2014



- co-variations in protein families ⇒ contact prediction

  Weigt et al. PNAS 2009; Morcos et al. PNAS 2011
  Marks et al PLoS ONE 2012; Sulkowska et al PNAS 2012

- diversity of antibody repertoires in the immune system

  Mora Walczak Callan Bialek PNAS 2010

- DNA motifs of transcription factor binding sites

  Santolini Mora Hakim Plos ONE 2014

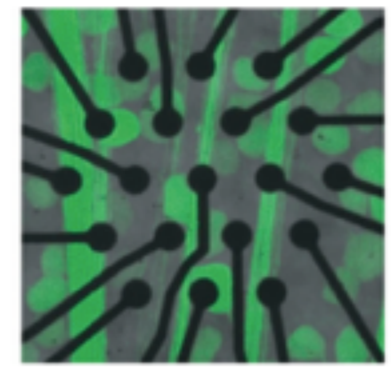
- collective behaviour of mice

  Shemesh et al. eLife 2013

maximum entropy in biology

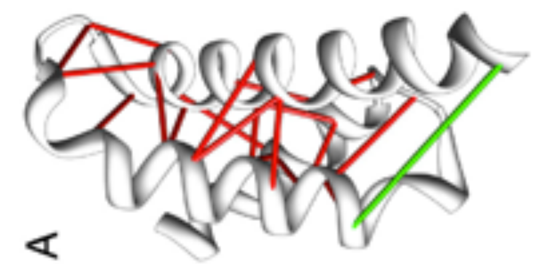
- collective activity of neural populations

  Schneidman et al. Nature 2006
  Shlens et al J. Neuroscience 2006
  Tang et al J. Neuroscience 2008
  Tkacik et al PLoS CP 2014



- co-variations in protein families ⇒ contact prediction

  Weigt et al. PNAS 2009; Morcos et al. PNAS 2011
  Marks et al PLoS ONE 2012; Sulkowska et al PNAS 2012



- diversity of antibody repertoires in the immune system

  Mora Walczak Callan Bialek PNAS 2010

- DNA motifs of transcription factor binding sites

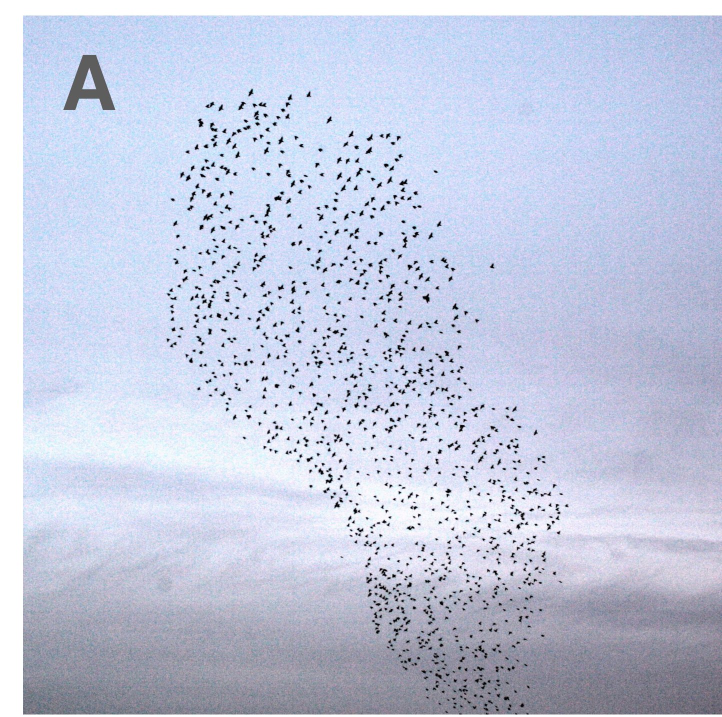
  Santolini Mora Hakim Plos ONE 2014

- collective behaviour of mice

  Shemesh et al. eLife 2013

- collective behaviour of bird flocks

  Bialek et al. PNAS 2012; Cavagna et al. PRE 2014
  Bialek et al. PNAS 2014

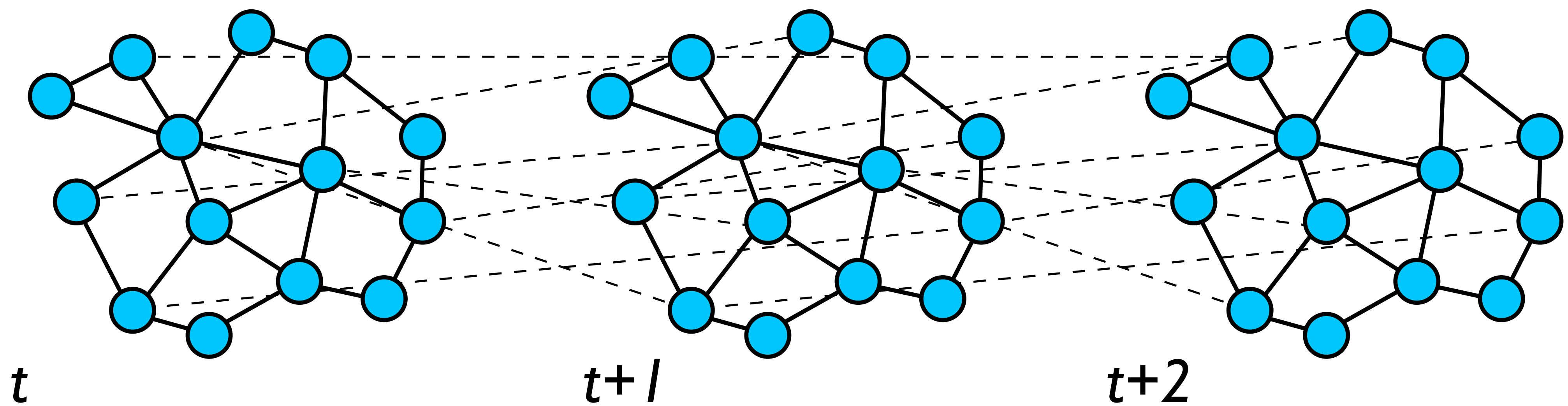


dynamical maximum entropy

- maximum entropy gives a “steady-state” picture.
- what about the dynamics?
- *ad hoc* dynamics such as Glauber, Metropolis may be wrong

dynamical maximum entropy

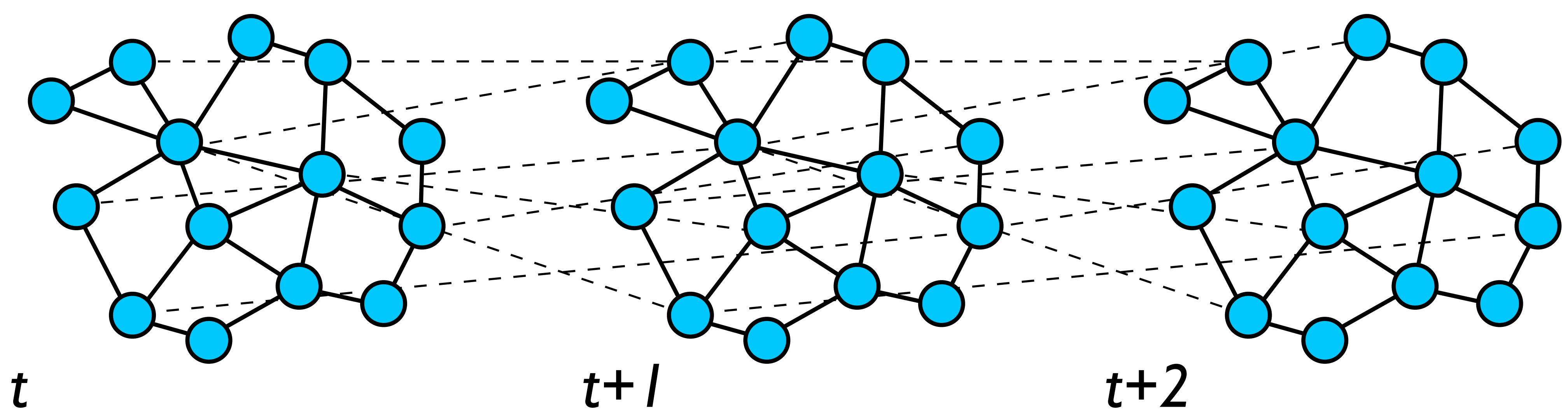
- maximum entropy gives a "steady-state" picture.
- what about the dynamics?
- *ad hoc* dynamics such as Glauber, Metropolis may be wrong
- (a) solution: maximum entropy over trajectories $P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^T)$



constraints on cross-time correlations, e.g. $\langle \sigma_i^t \sigma_j^{t'} \rangle$

dynamical maximum entropy

- maximum entropy gives a "steady-state" picture.
- what about the dynamics?
- *ad hoc* dynamics such as Glauber, Metropolis may be wrong
- (a) solution: maximum entropy over trajectories $P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^T)$

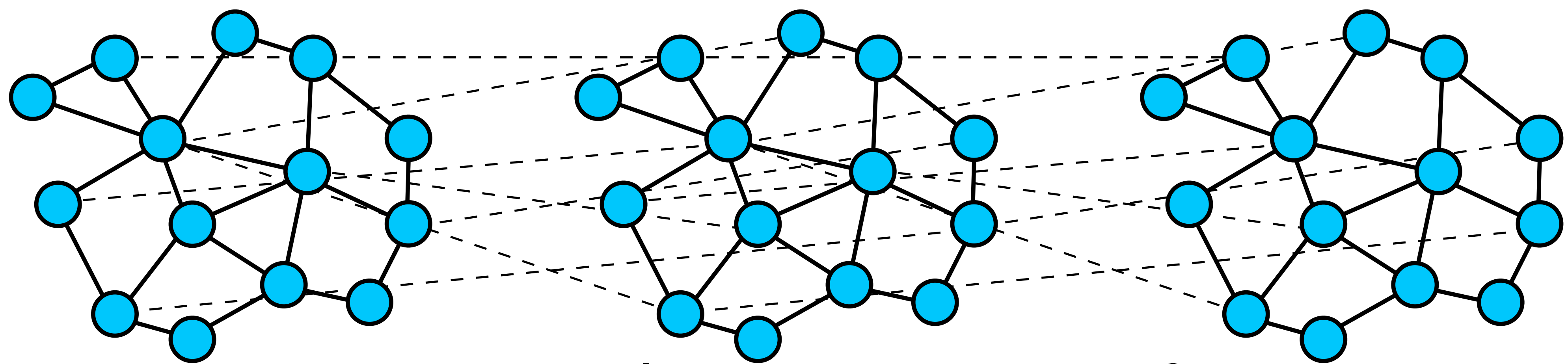


constraints on cross-time correlations, e.g. $\langle \sigma_i^t \sigma_j^{t'} \rangle$

$$\Longrightarrow \qquad P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^T) = \frac{1}{Z} \exp \left(\overbrace{\sum_{i,j,t,t'} J_{ij}^{t-t'} \sigma_i^t \sigma_j^{t'}}^{-\mathcal{A} \text{ “action”}} \right)$$

dynamical maximum entropy

- maximum entropy gives a "steady-state" picture.
- what about the dynamics?
- *ad hoc* dynamics such as Glauber, Metropolis may be wrong
- (a) solution: maximum entropy over trajectories $P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^T)$



constraints on cross-time correlations, e.g. $\langle \sigma_i^t \sigma_j^{t'} \rangle$

$$\Longrightarrow \qquad P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^T) = \frac{1}{Z} \exp \Bigg(\overbrace{\sum_{i,j,t,t'} J_{ij}^{t-t'} \sigma_i^t \sigma_j^{t'}}^{-\mathcal{A} \text{ “action”}} \Bigg)$$

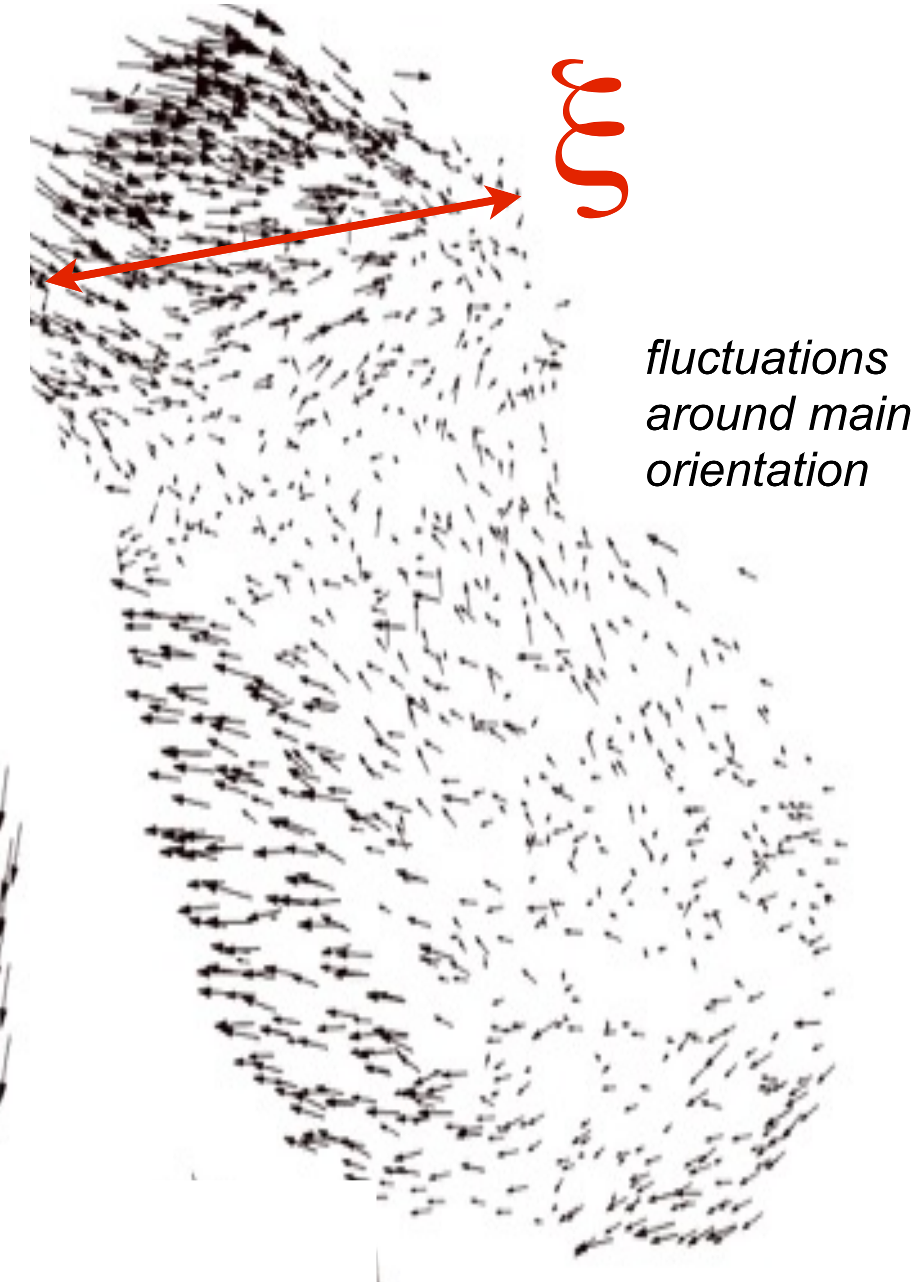
- not the same as:

$$P(\sigma_{i,t} | \{\sigma_{j,t'}\}_{t'<t}) = \frac{1}{Z(\{\sigma_{j,t'}\}_{t'<t})} \exp \left[h_i \sigma_{i,t} + \sum_{j,t'<t} J_{ij}^{t-t'} \sigma_{i,t} \sigma_{j,t'} \right]$$

example 1: flocks of birds

example 1: flocks of birds

aligned collective motion



strong polarization

$$\phi = \left| \frac{1}{N} \sum_i \frac{\vec{v}_i}{|\vec{v}_i|} \right| \sim 0.95$$

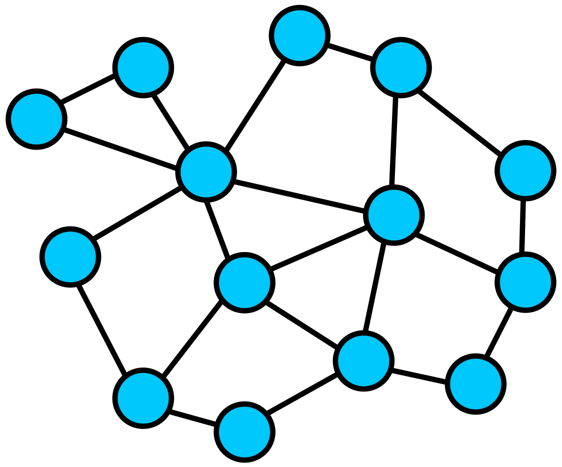
domains

a maximum entropy model for birds

- velocity of bird $\vec{v}_i, \quad \vec{s}_i = \vec{v}_i / \|\vec{v}_i\|$

a maximum entropy model for birds

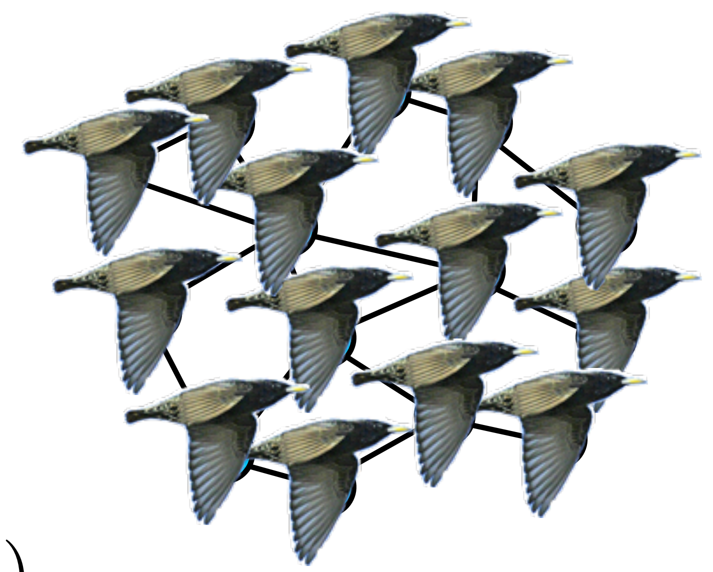
- velocity of bird $\vec{v}_i, \quad \vec{s}_i = \vec{v}_i / \|\vec{v}_i\|$
- constrain correlation functions $C_{ij} = \langle \vec{s}_i \vec{s}_j \rangle$



$$P(\vec{s}_1, \ldots, \vec{s}_N) = \frac{1}{Z} \exp\left(\sum_{ij} J_{ij} \vec{s}_i \vec{s}_j \right) = \frac{1}{Z} \exp(-H)$$

(Heisenberg model on lattice)

a maximum entropy model for birds

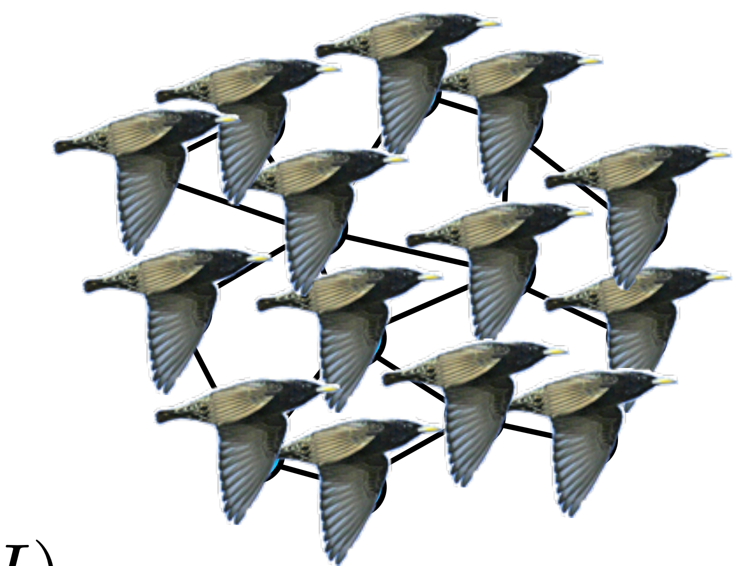


- velocity of bird $\vec{v}_i, \quad \vec{s}_i = \vec{v}_i / ||\vec{v}_i||$
- constrain correlation functions $C_{ij} = \langle \vec{s}_i \vec{s}_j \rangle$

$$P(\vec{s}_1, \ldots, \vec{s}_N) = \frac{1}{Z} \exp\left(\sum_{ij} J_{ij} \vec{s}_i \vec{s}_j \right) = \frac{1}{Z} \exp(-H)$$

(Heisenberg model on lattice)

a maximum entropy model for birds



- velocity of bird $\vec{v}_i, \quad \vec{s}_i = \vec{v}_i / \|\vec{v}_i\|$
- constrain correlation functions $C_{ij} = \langle \vec{s}_i \vec{s}_j \rangle$

$$P(\vec{s}_1, \ldots, \vec{s}_N) = \frac{1}{Z} \exp\left(\sum_{ij} J_{ij} \vec{s}_i \vec{s}_j \right) = \frac{1}{Z} \exp(-H)$$

(Heisenberg model on lattice)

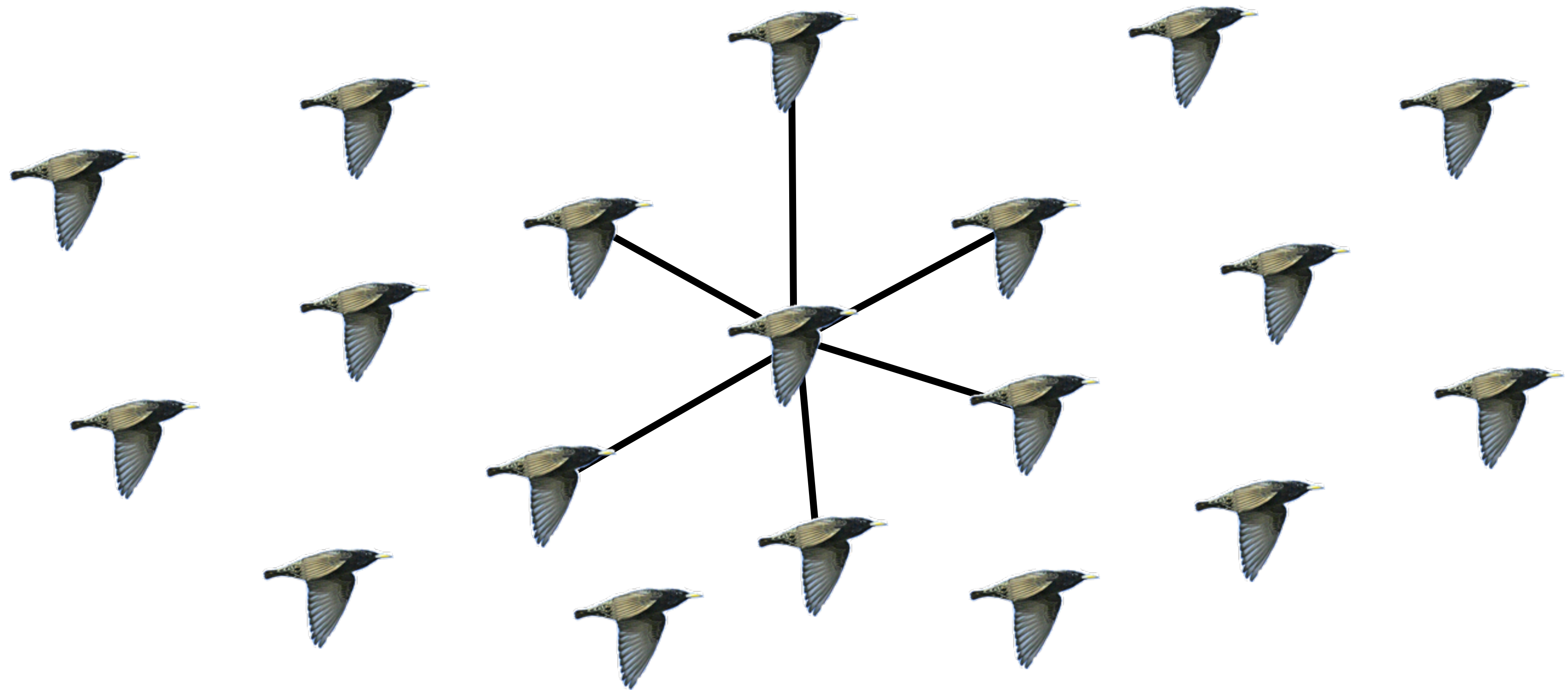
- derives from Langevin eqn, equilavent to "social" model, similar to Vicsek's

$$\frac{d\vec{s}_\mathrm{i}}{dt} = -\frac{\partial H}{\partial \vec{s}_\mathrm{i}} + \vec{\eta}_\mathrm{i}(t) = \sum_{\mathrm{j}=1}^{N} J_{\mathrm{ij}} \vec{s}_\mathrm{j} + \vec{\eta}_\mathrm{i}(t)$$

alignment noise

(does not mean that's the only possible dynamics, or the true one)

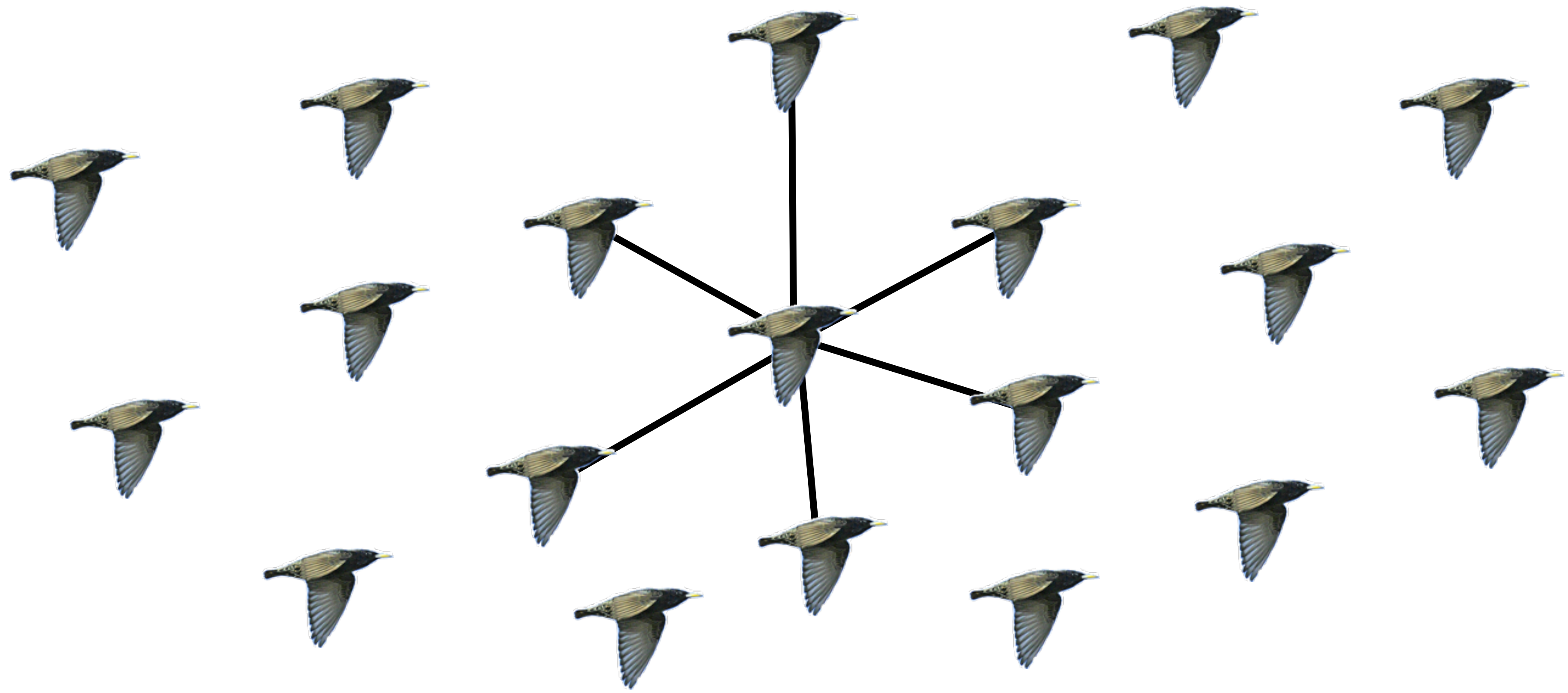
parametrization



$$J_{ij} = \begin{cases} J & \text{if j is one i's } n_c \text{ first neighbors} \\ 0 & \text{otherwise} \end{cases}$$

then symmetrized

parametrization



$$J_{ij} = \begin{cases} J & \text{if j is one i's } n_c \text{ first neighbors} \\ 0 & \text{otherwise} \end{cases}$$

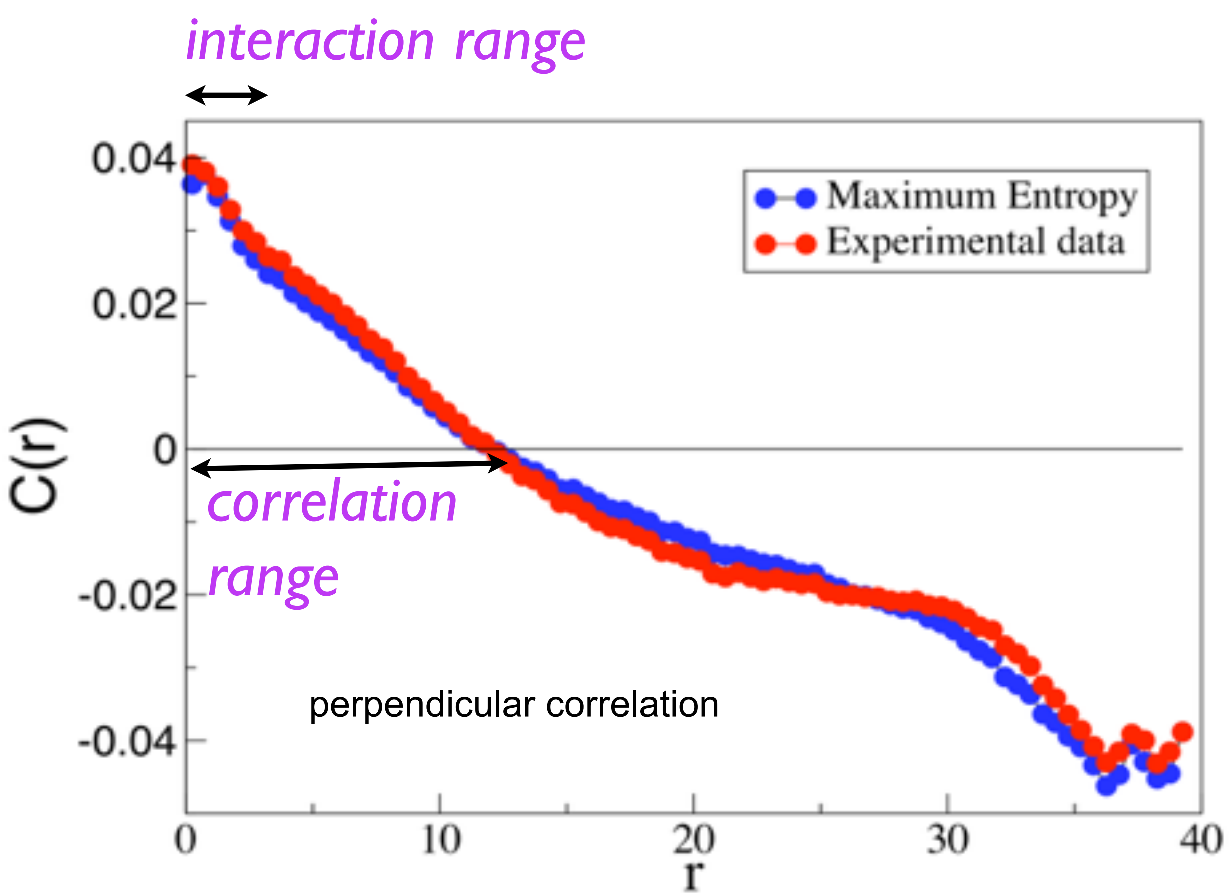
then symmetrized

Equivalent to maximum entropy with constraint on

$$C_{\text{int}} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_c} \sum_{j \in V(i)} \langle \vec{s}_i \vec{s}_j \rangle$$

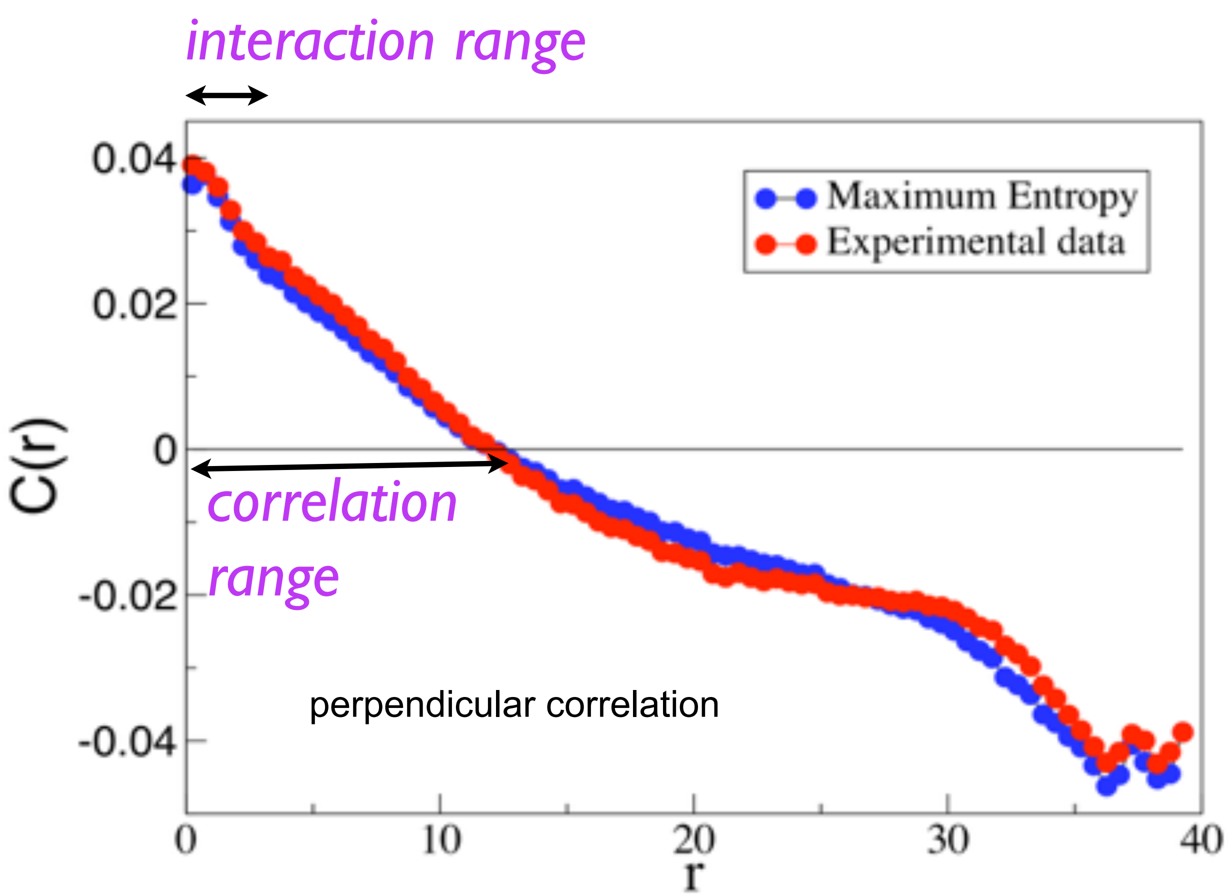
single snapshot — spatial averaging instead of ensemble averaging

predicting correlation functions

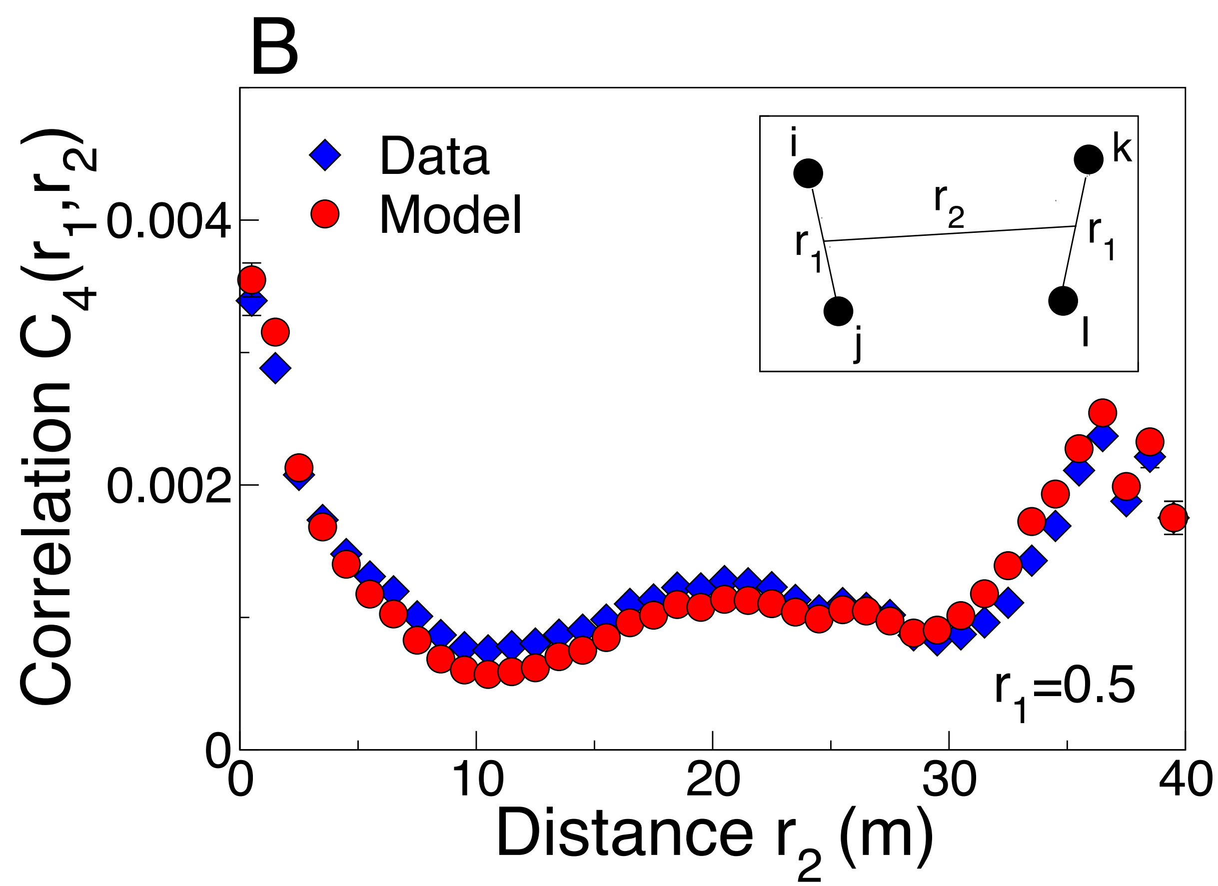


long-range order from local interactions

predicting correlation functions



4-bird correlation function

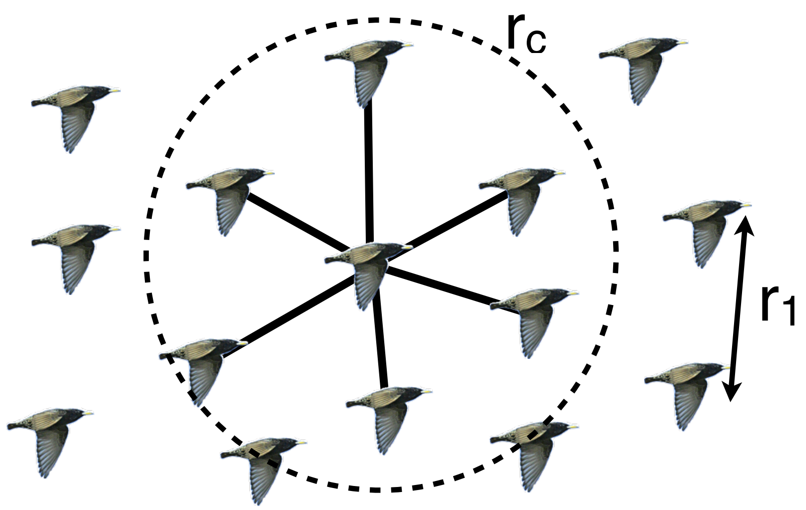


long-range order from local interactions

interaction range

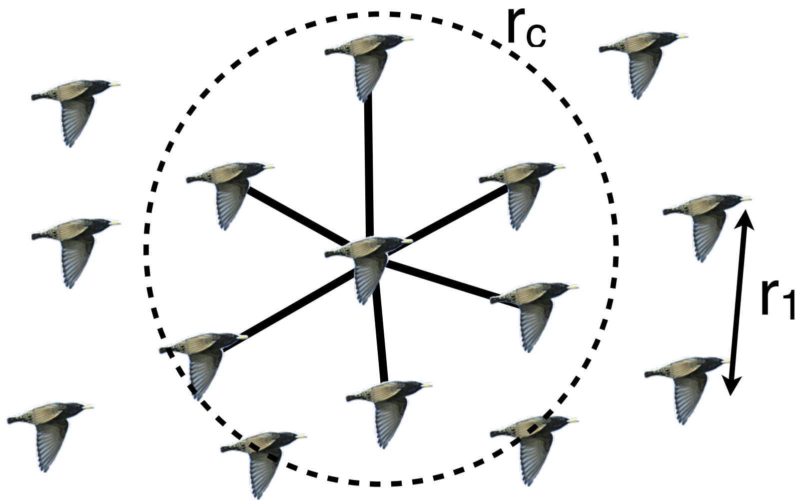
metric or topological ?

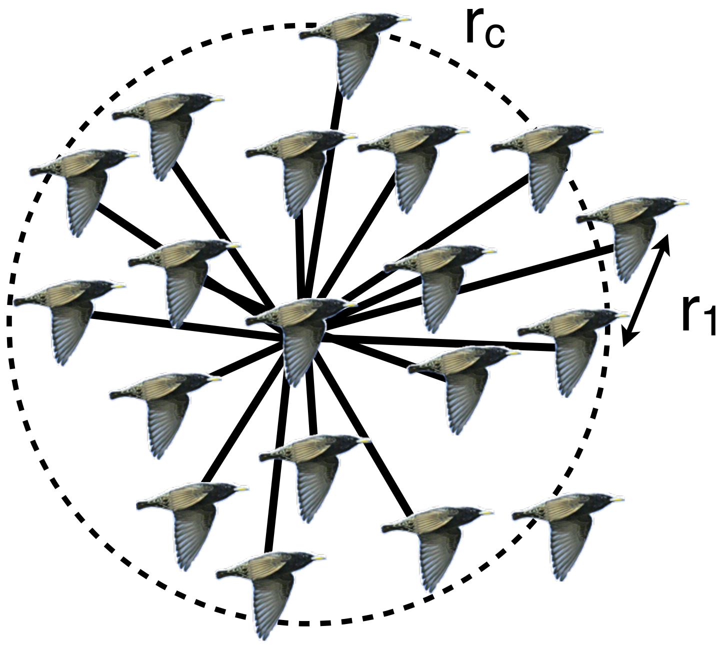
interaction range



metric or topological ?

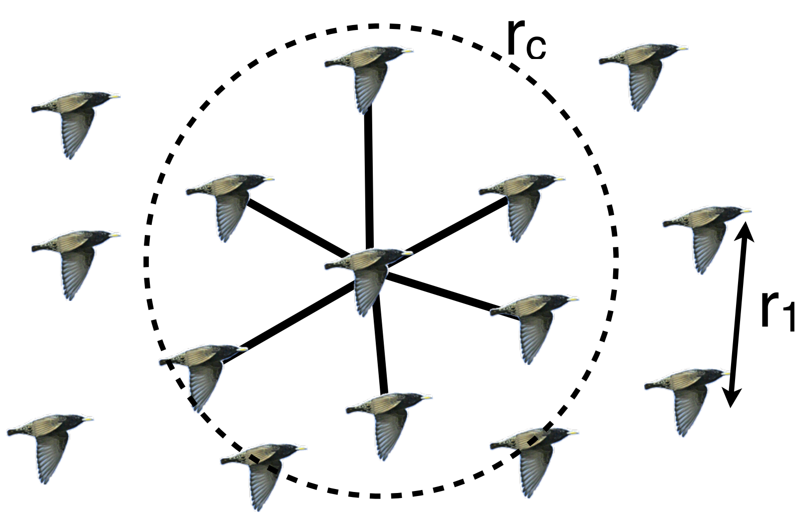
interaction range

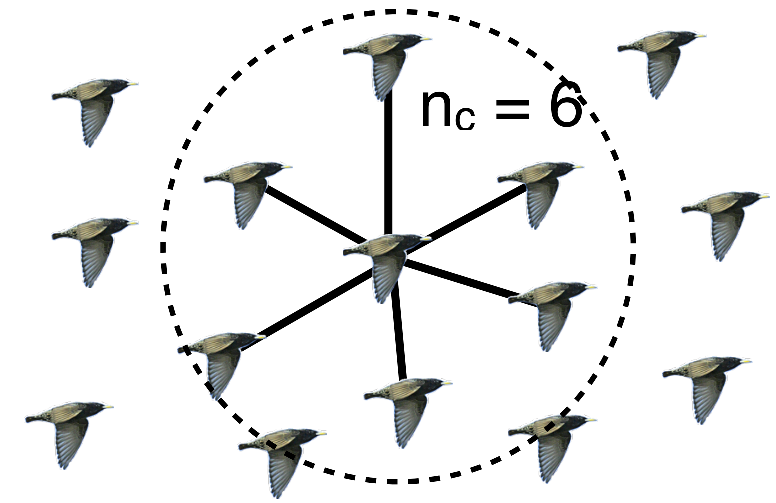




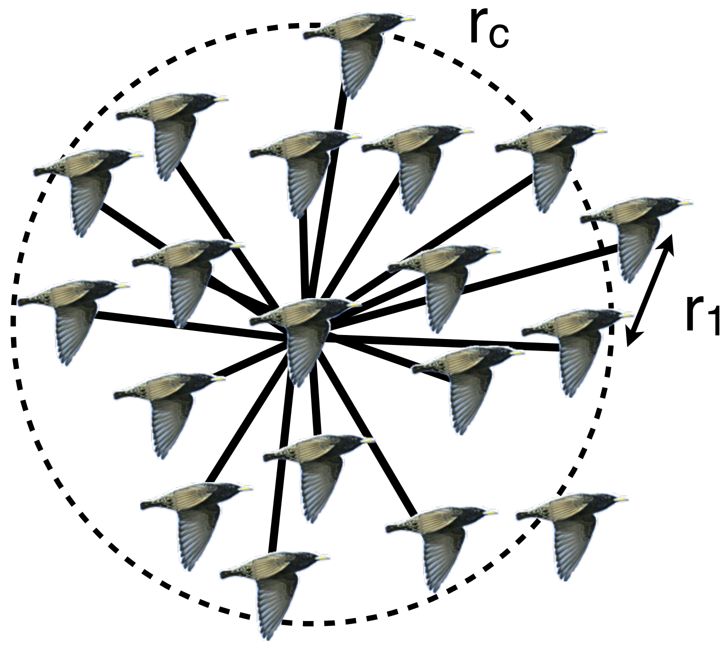
metric or topological ?

interaction range

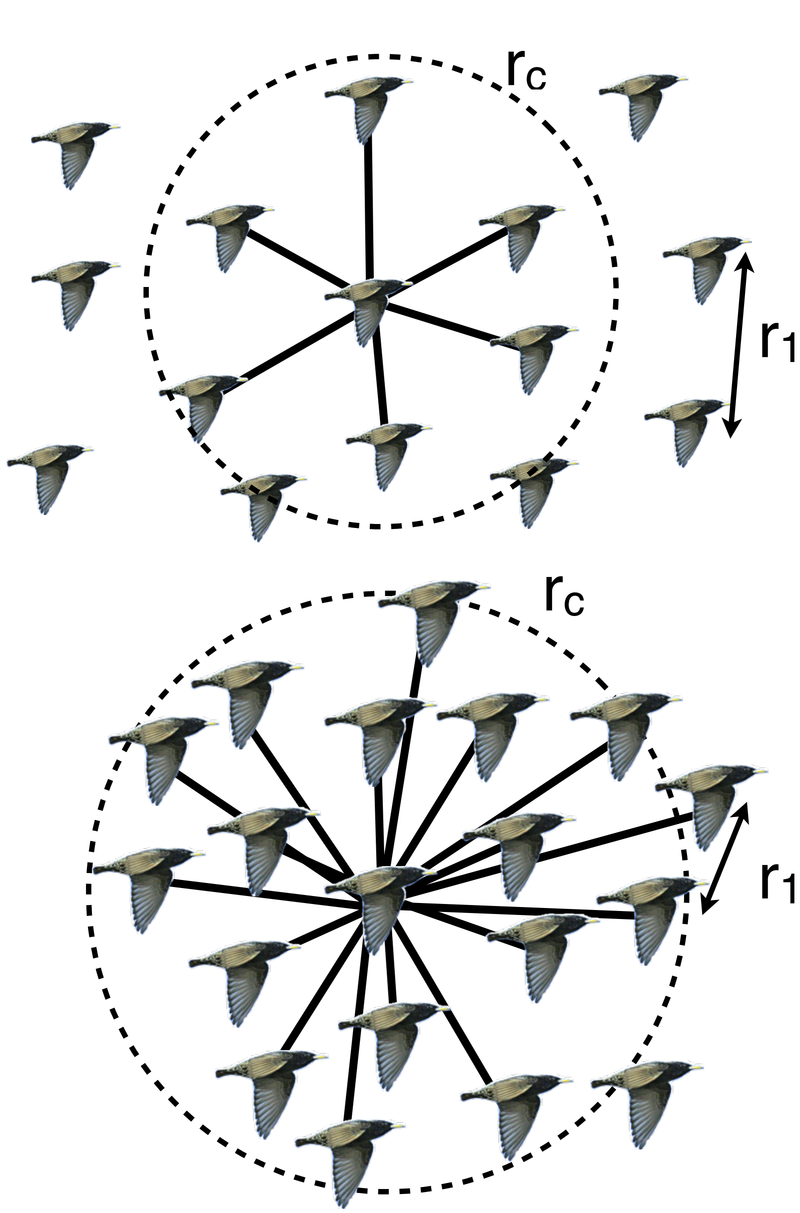




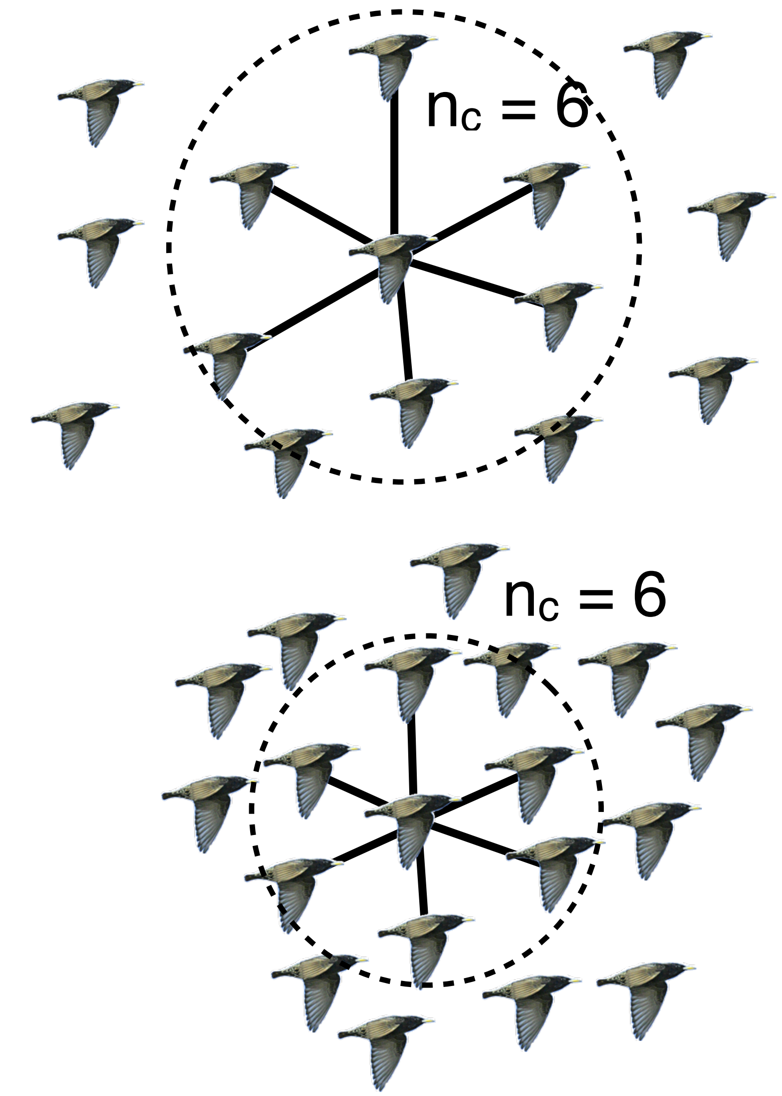
metric or topological ?



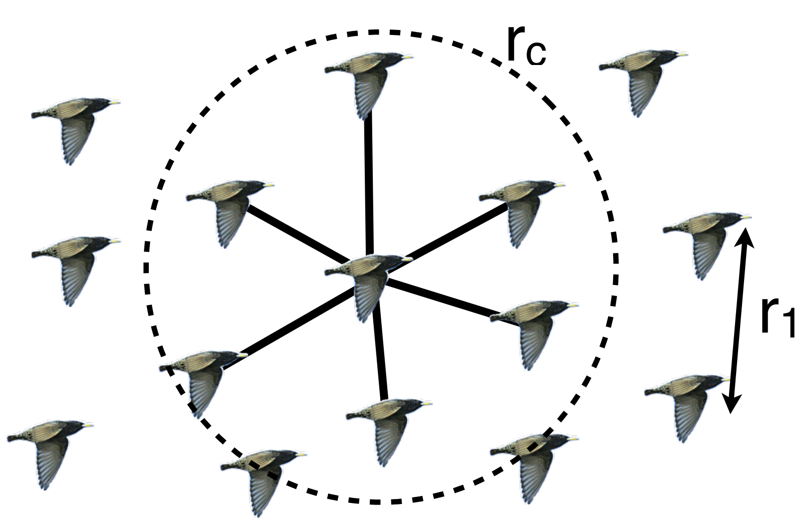
interaction range

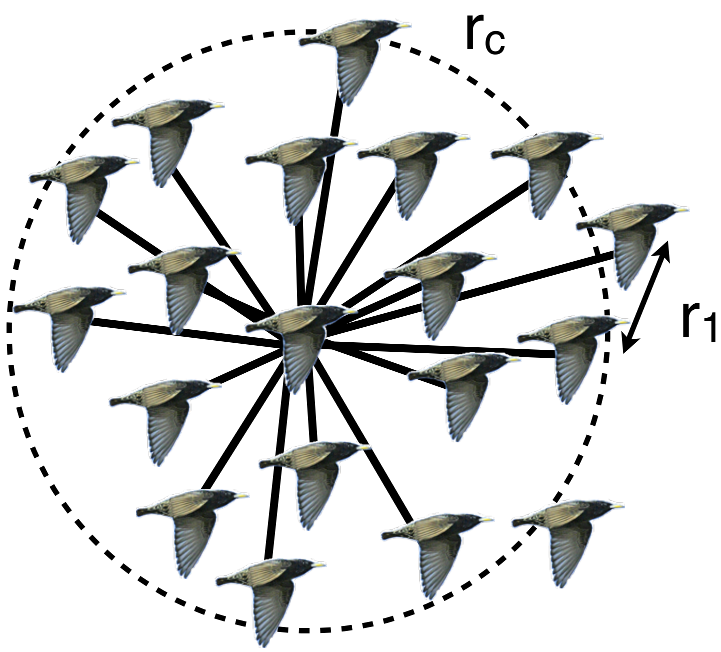


metric or topological ?

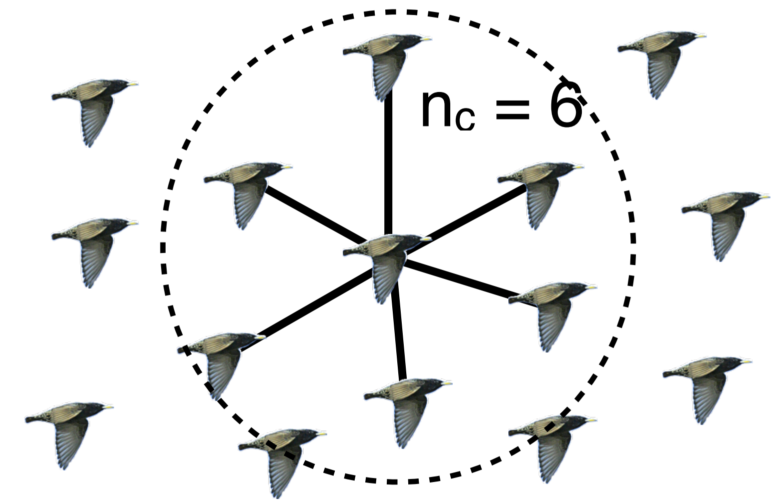


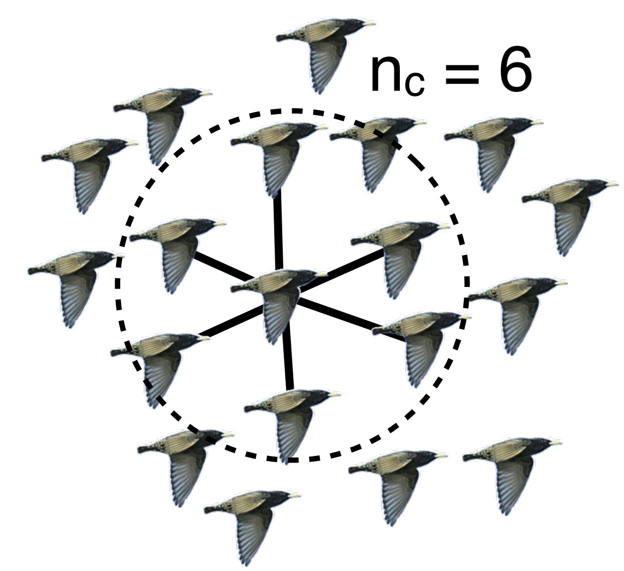
interaction range

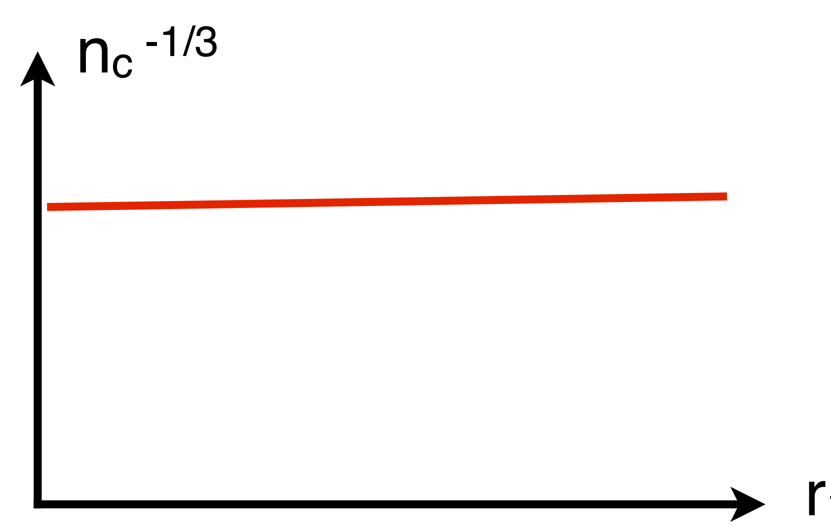




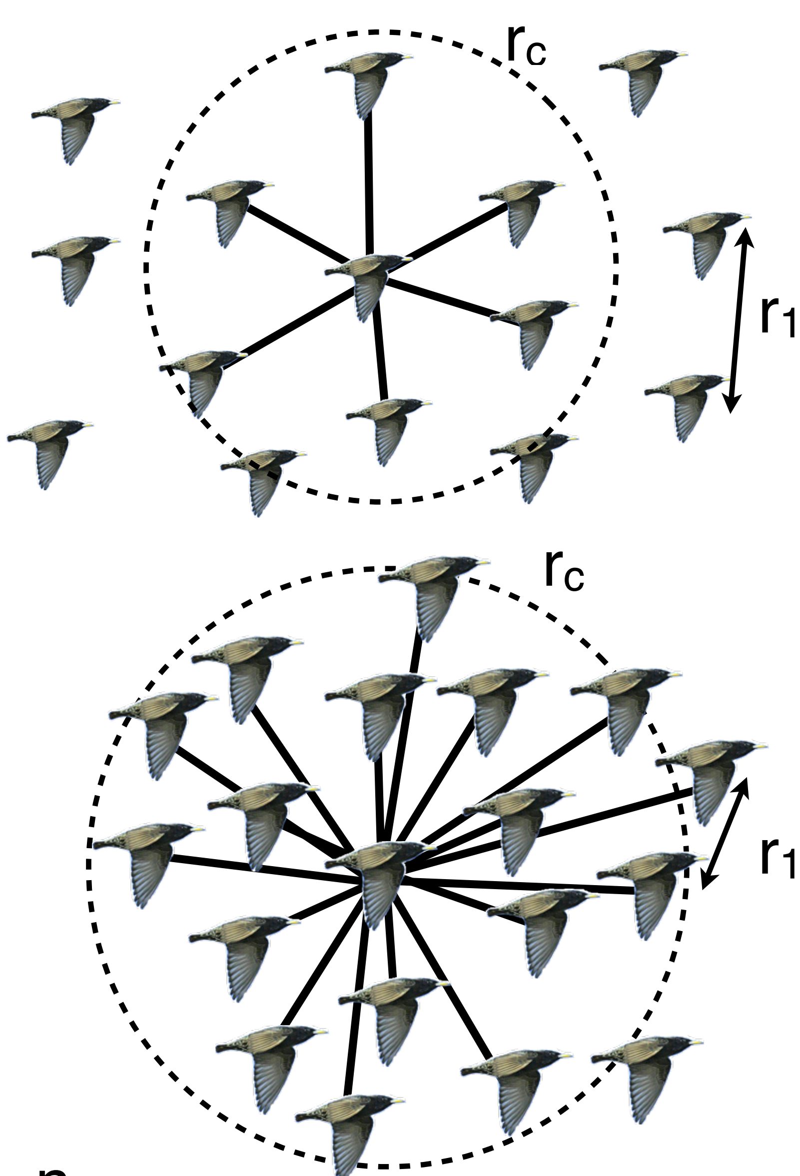
metric or topological ?



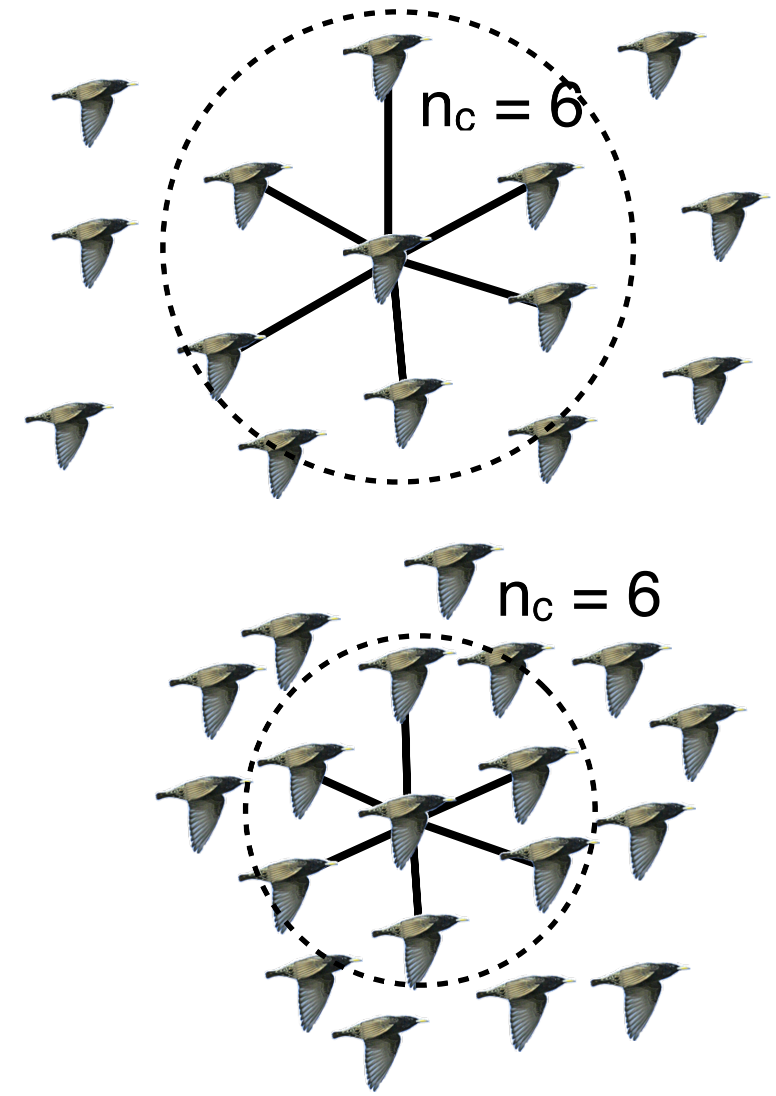




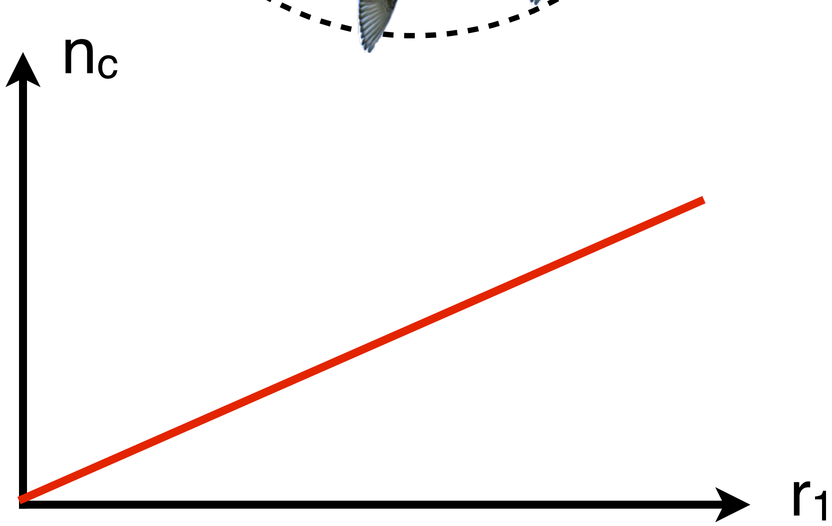
interaction range

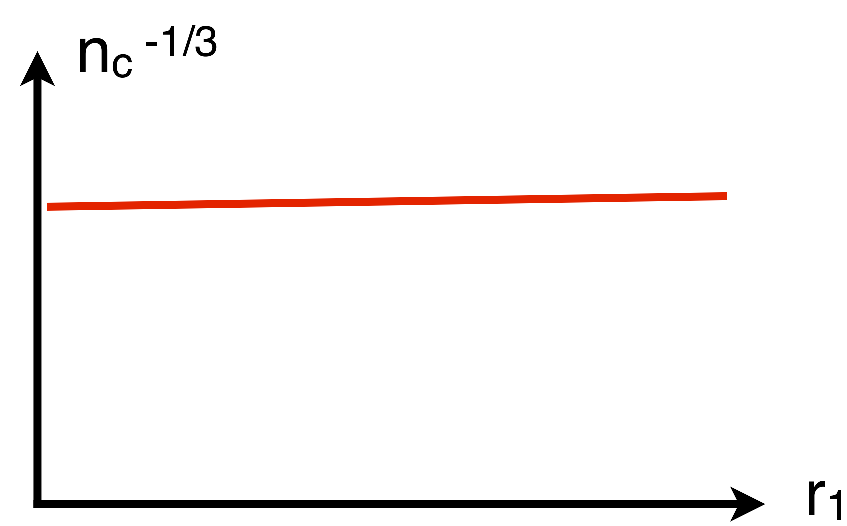


metric or topological ?



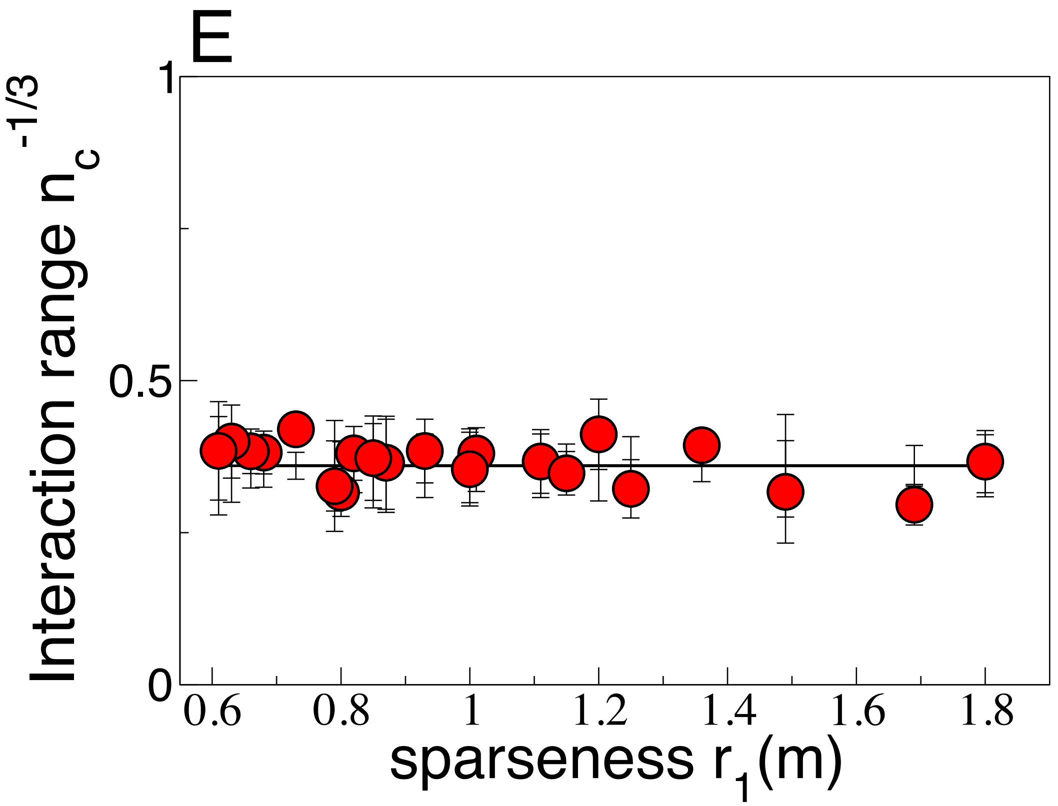
$$n_c \sim (r_c / r_1)^3$$





answer:

interaction is
topological not metric



Bialek et al PNAS 2012

answer:

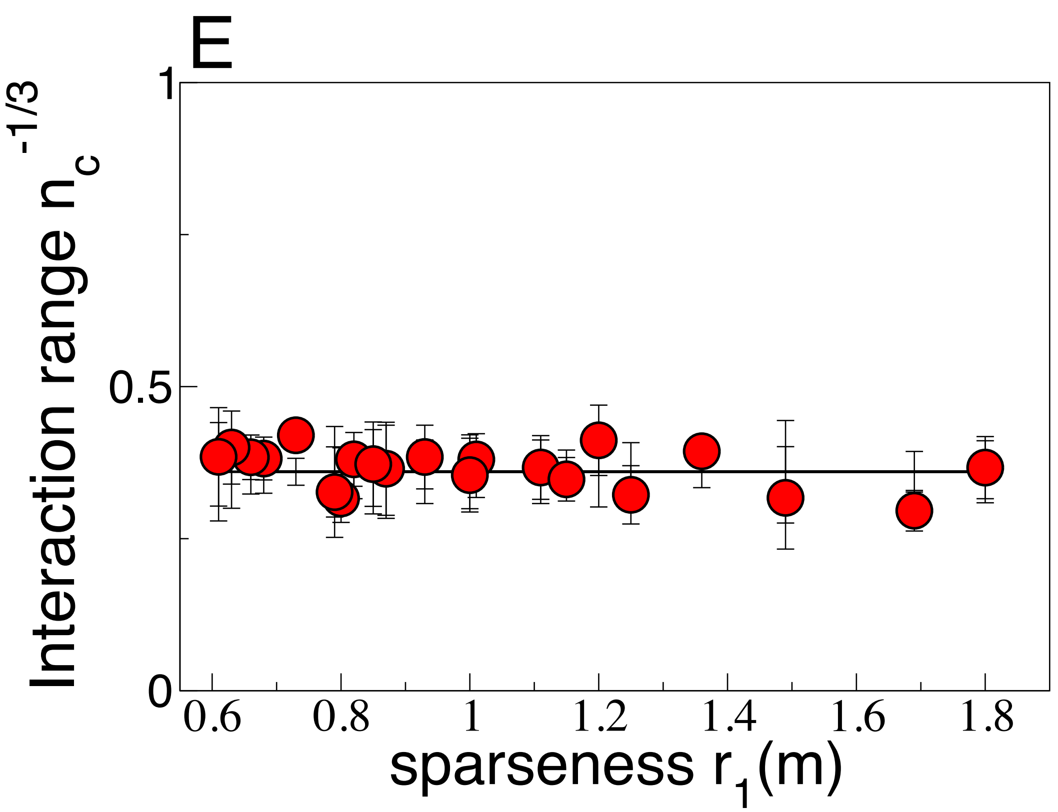
interaction is
topological not metric

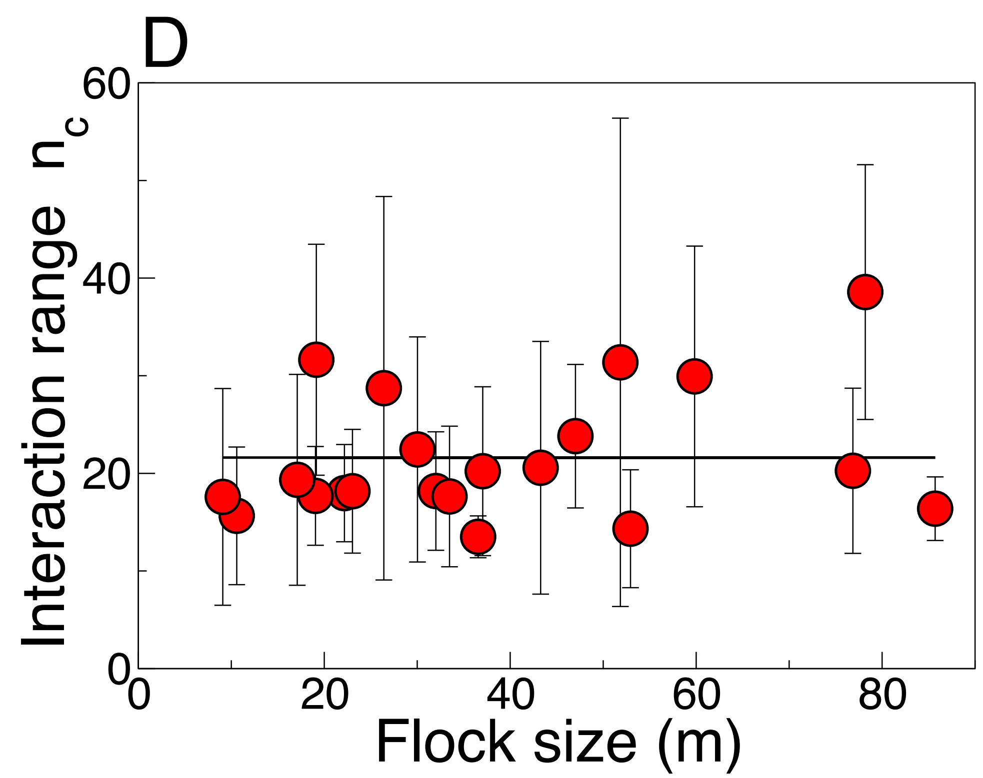
$n_c \sim 21$

does not depend on

flock density

flock size

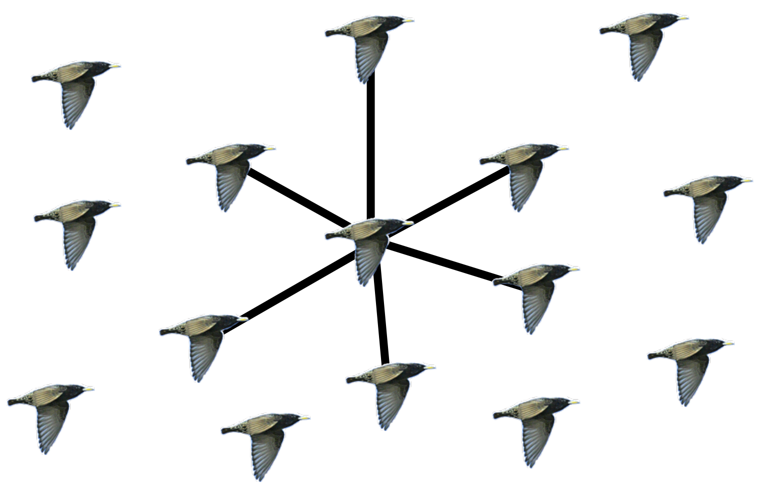




Bialek et al PNAS 2012

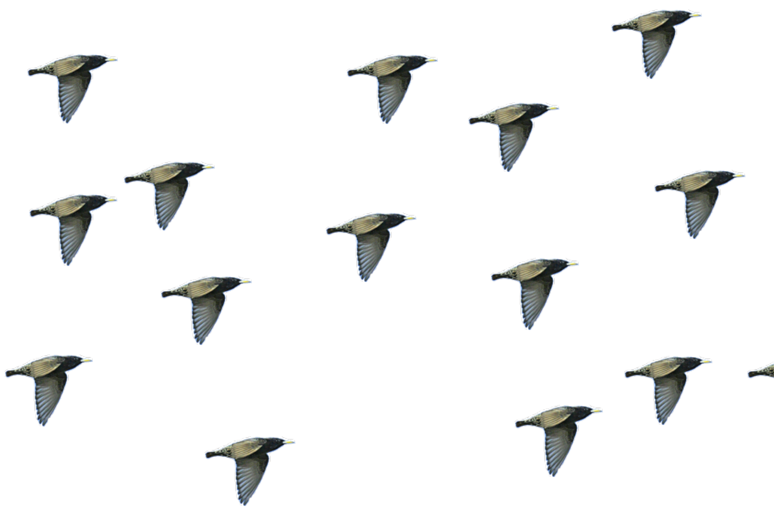
dynamics (may) matter

- we've assumed that neighborhoods are fixed
- but birds may exchange neighbors fast



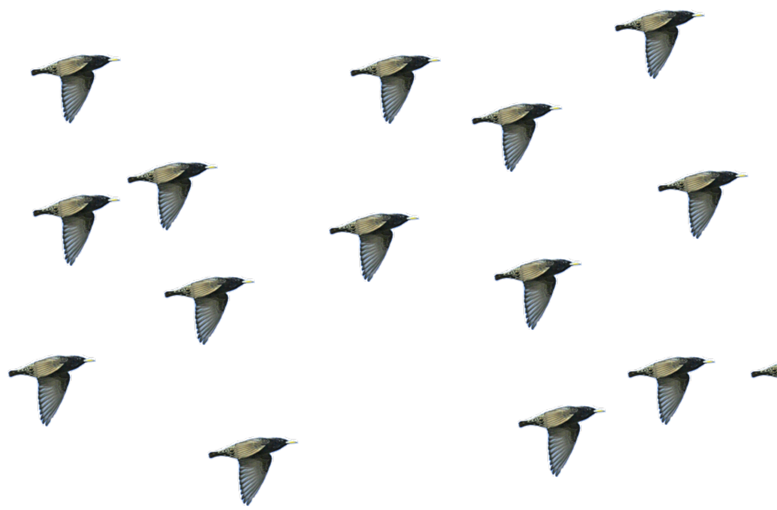
dynamics (may) matter

- we’ve assumed that neighborhoods are fixed
- but birds may exchange neighbors fast



dynamics (may) matter

- we’ve assumed that neighborhoods are fixed
- but birds may exchange neighbors fast
- the *effective* number of interaction partners could be larger than the *instantaneous* one.



dynamics (on bird orientations)

- constrain $\langle s_i^t s_j^t \rangle$ *and* $\langle s_i^t s_j^{t+1} \rangle$

$$P(s^1, \ldots, s^T) = \frac{1}{\hat{Z}} \exp\left(-\mathcal{A}\right)$$

"action" $$\mathcal{A} = -\frac{1}{2} \sum_t \sum_{i \neq j} \left(J_{ij;t}^{(1)} s_i^t s_j^t + J_{ij;t}^{(2)} s_i^{t+1} s_j^t \right)$$

dynamics (on bird orientations)

- constrain $\langle s_i^t s_j^t \rangle$ *and* $\langle s_i^t s_j^{t+1} \rangle$

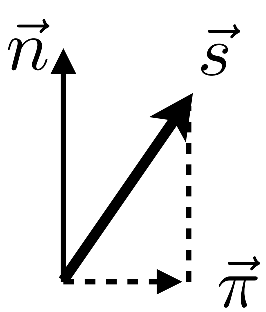
$$P(s^1, \ldots, s^T) = \frac{1}{\hat{Z}} \exp\left(-\mathcal{A}\right)$$

"action" $$\mathcal{A} = -\frac{1}{2} \sum_t \sum_{i \neq j} \left(J_{ij;t}^{(1)} s_i^t s_j^t + J_{ij;t}^{(2)} s_i^{t+1} s_j^t \right)$$

- in spin-wave approximation, equivalent to "collective random walk"

$$\pi_i(t+1) = \sum_j M_{ij}(t) \pi_j(t) + \epsilon_i(t)$$

$$\langle \epsilon(t)^\dagger \epsilon(t') \rangle = 2(d-1) A(t)^{-1} \delta_{t,t'}$$



A and M functions of $J^{(1)}$ and $J^{(2)}$

dynamics (on bird orientations)

- constrain $\langle s_i^t s_j^t \rangle$ *and* $\langle s_i^t s_j^{t+1} \rangle$

$$P(s^1, \ldots, s^T) = \frac{1}{\hat{Z}} \exp\left(-\mathcal{A}\right)$$

"action" $$\mathcal{A} = -\frac{1}{2} \sum_t \sum_{i \neq j} \left(J_{ij;t}^{(1)} s_i^t s_j^t + J_{ij;t}^{(2)} s_i^{t+1} s_j^t \right)$$

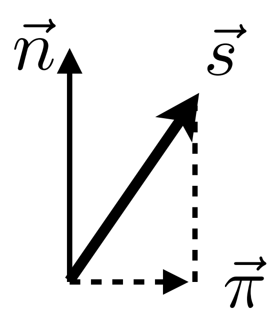
- in spin-wave approximation, equivalent to "collective random walk"

alignment strength

$$\pi_i(t+1) = (1 - J\delta t\, n_c)\pi_i(t) + J\delta t\, n_{ij}(t) + \epsilon_i(t)$$

$$\langle \epsilon_i(t) \epsilon_j(t') \rangle = 2(d-1)\delta t T \delta_{ij} \delta_{tt'}$$

temperature



Langevin equation

inferring out-of-equilibrium behavior

- infering J, n_c, and a third parameter, the "temperature" T

$$J\,n_c = \frac{1}{\delta t}\frac{C_{\rm int} - C_s + G_s - G_{\rm int}}{2C_{\rm int} - C'_{\rm int} - C_s}$$ and similar eq. for T

C_s^1	$(1/N)\sum_i(\pi_i^{t+1})^2$	$C_{\rm int}$	$(1/Nn_c)\sum_{ij} n_{ij}\pi_i^t\pi_j^t$
C_s	$(1/N)\sum_i(\pi_i^t)^2$	$C'_{\rm int}$	$(1/Nn_c^2)\sum_{ijk} n_{ij}n_{ik}\pi_j^t\pi_k^t$
G_s	$(1/N)\sum_i \pi_i^{t+1}\pi_i^t$	$G_{\rm int}$	$(1/Nn_c)\sum_{ij} n_{ij}\pi_i^{t+1}\pi_j^t$

inferring out-of-equilibrium behavior

- infering J, n_c, and a third parameter, the "temperature" T

$$J\,n_c = \frac{1}{\delta t}\,\frac{C_{\text{int}} - C_s + G_s - G_{\text{int}}}{2C_{\text{int}} - C'_{\text{int}} - C_s}$$ and similar eq. for T

C_s^1	$(1/N)\sum_i (\pi_i^{t+1})^2$	C_{int}	$(1/Nn_c)\sum_{ij} n_{ij}\pi_i^t\pi_j^t$
C_s	$(1/N)\sum_i (\pi_i^t)^2$	C'_{int}	$(1/Nn_c^2)\sum_{ijk} n_{ij}n_{ik}\pi_j^t\pi_k^t$
G_s	$(1/N)\sum_i \pi_i^{t+1}\pi_i^t$	G_{int}	$(1/Nn_c)\sum_{ij} n_{ij}\pi_i^{t+1}\pi_j^t$

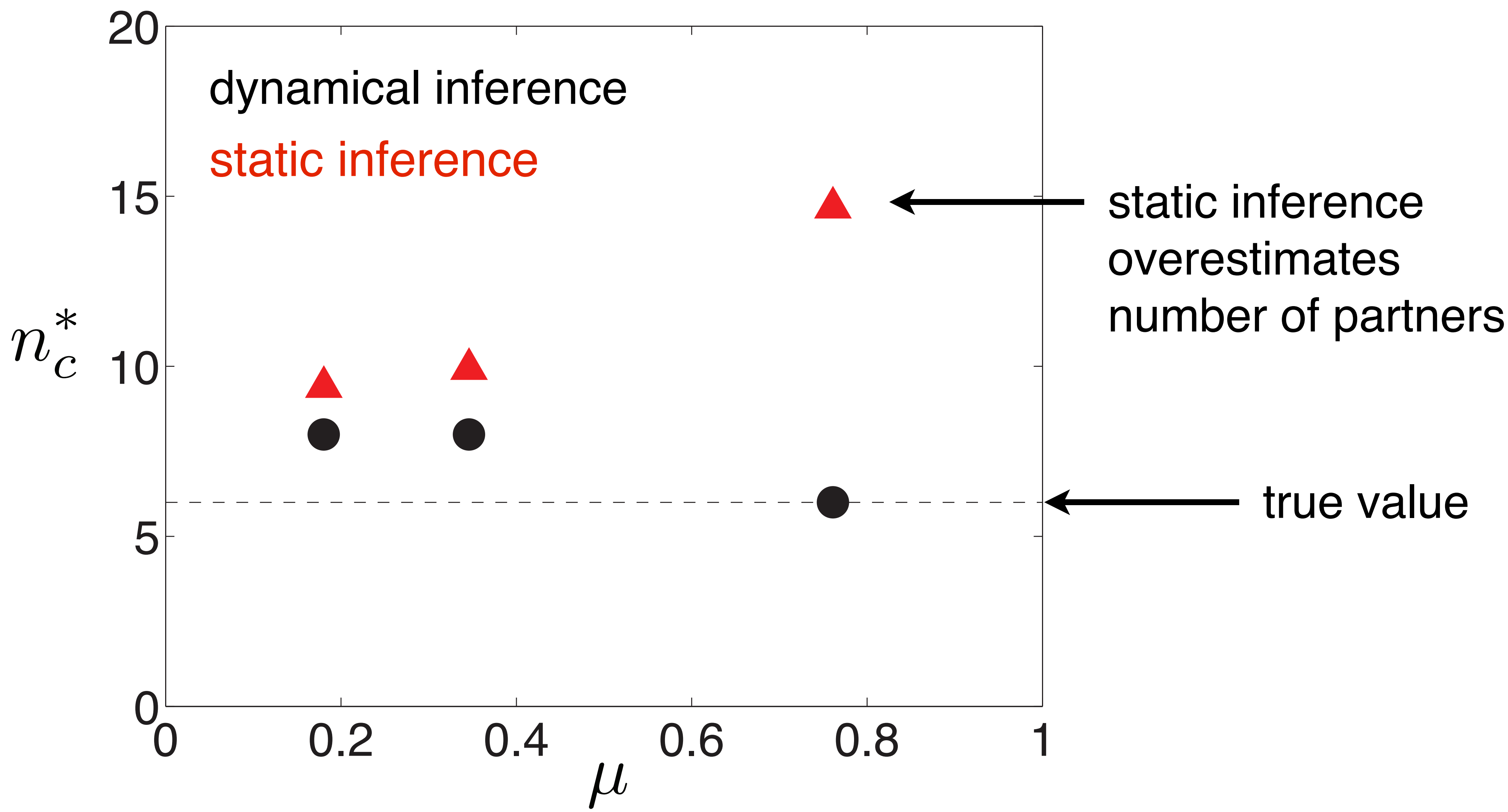
- if equilibrium – slowly evolving and symmetric n_{ij} – then

one recovers the same result as the Heisenberg model, with J ← J / T

$$P(\vec{s}_1, \ldots, \vec{s}_N) = \frac{1}{Z}\exp\left(\frac{J}{T}\sum_{ij} n_{ij}\vec{s}_i\vec{s}_j\right)$$

test on simulated data

- simulation of 2D topological model with Voronoi neighbors
- μ is a parameter quantifying how fast birds change neighbors

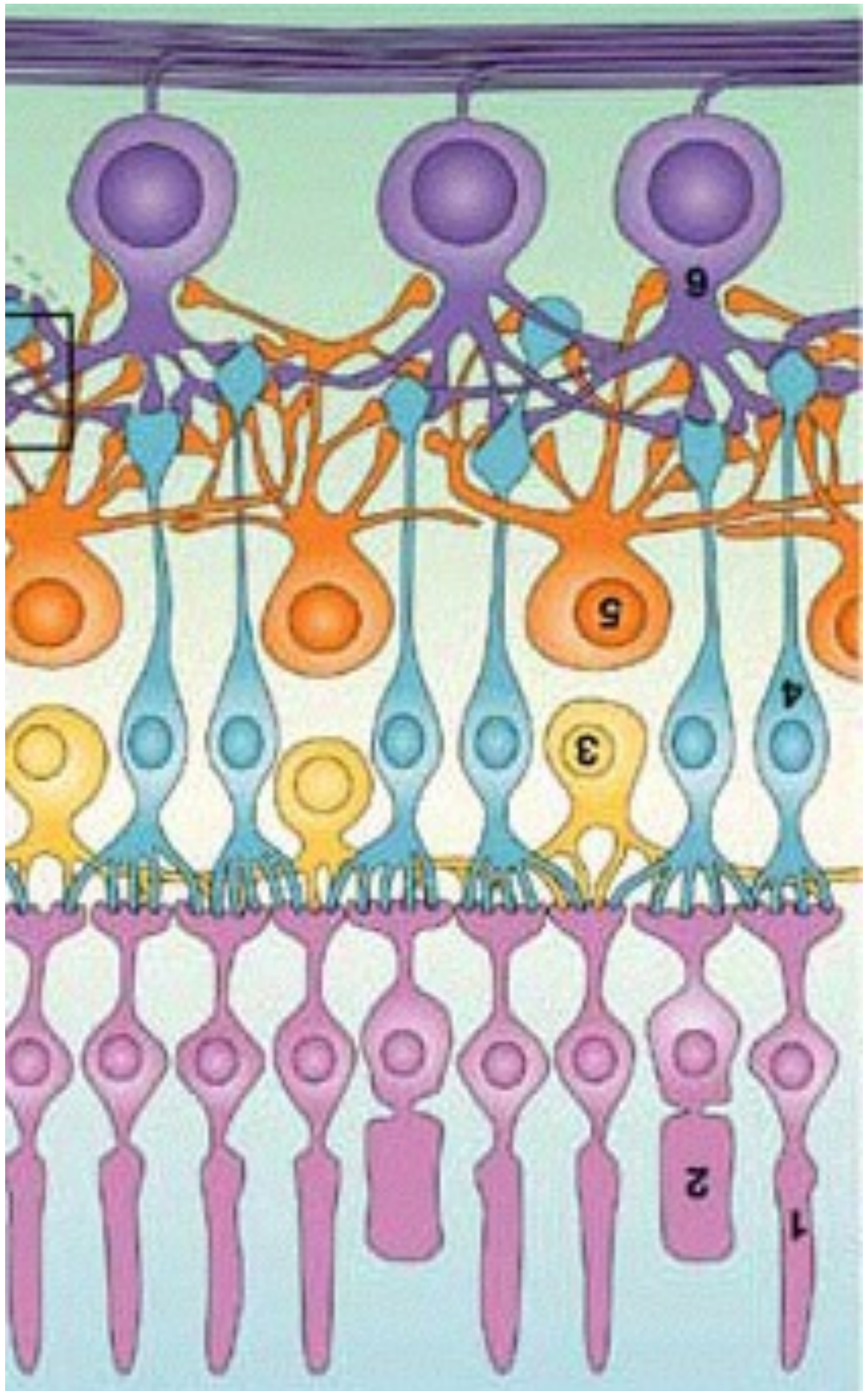


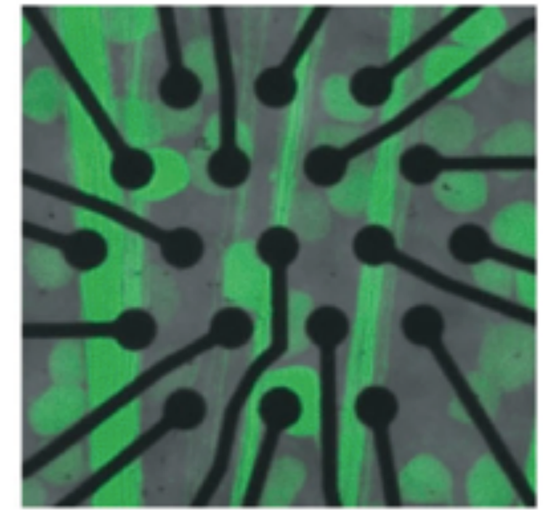
- at large μ,

dynamical maximum entropy works, static maximum entropy doesn't

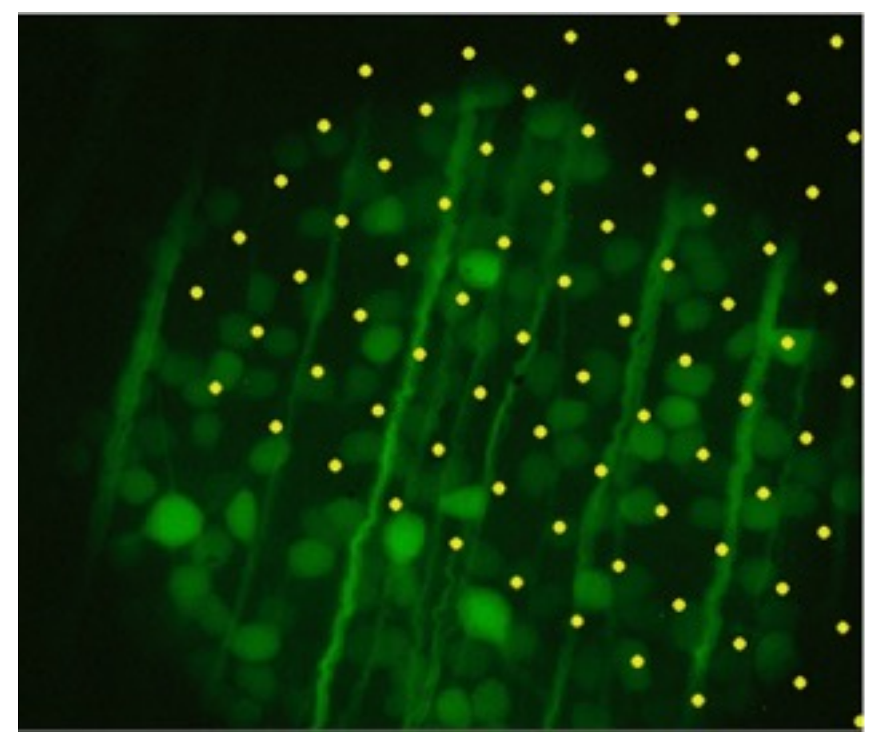
Cavagna et al PRE 2014

the retina





multielectrode array recordings

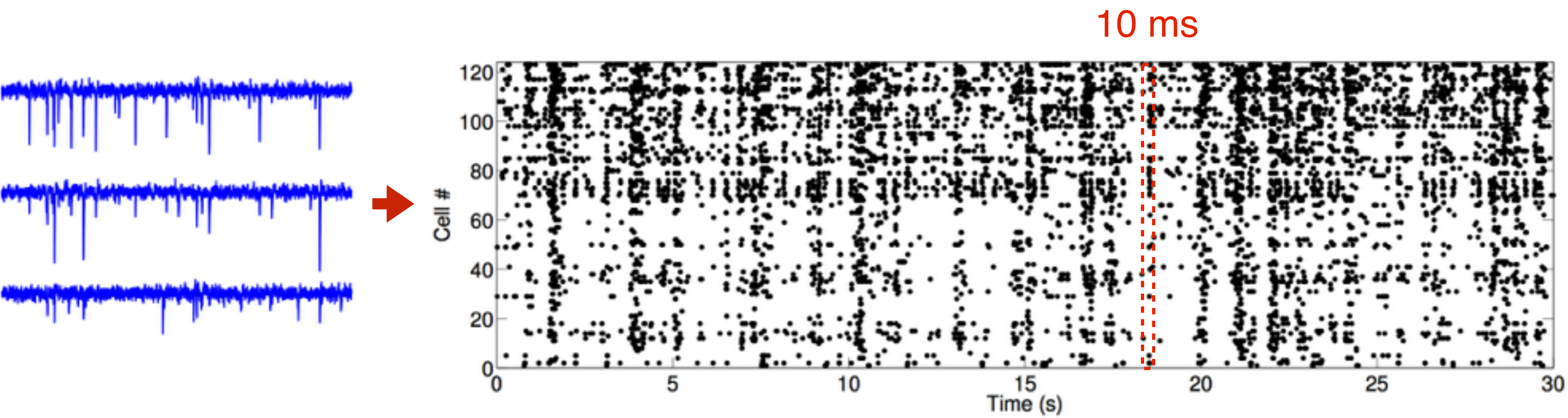


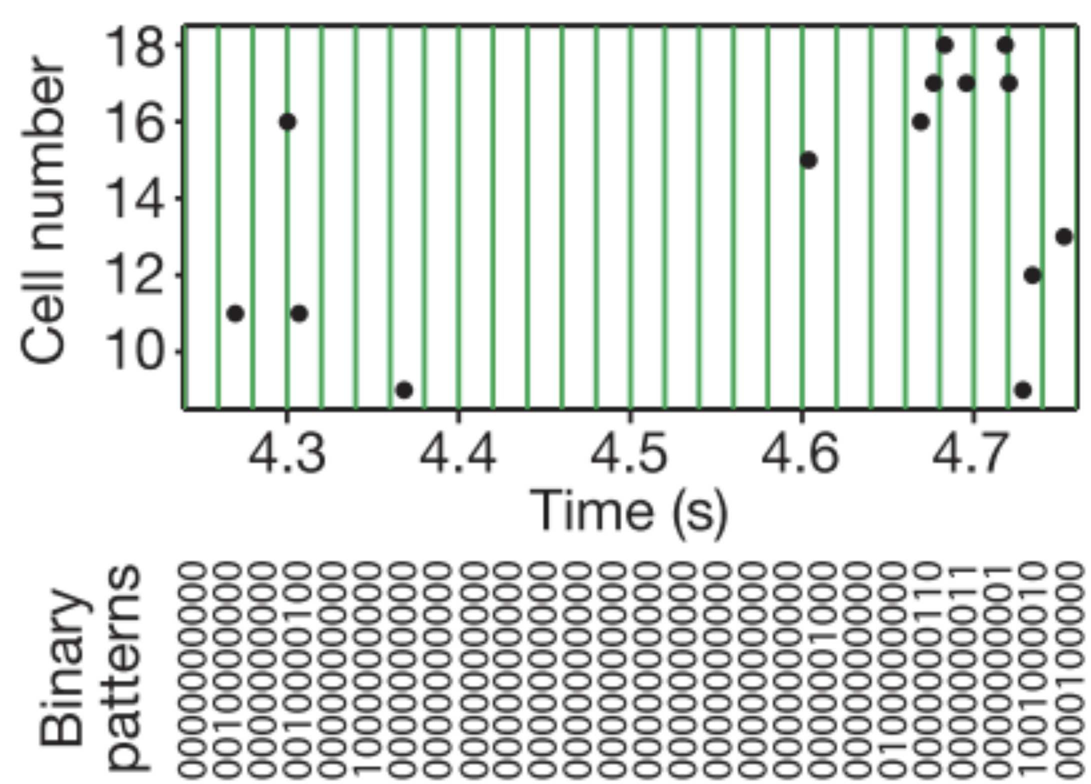
the stimulus

the stimulus

binary neurons

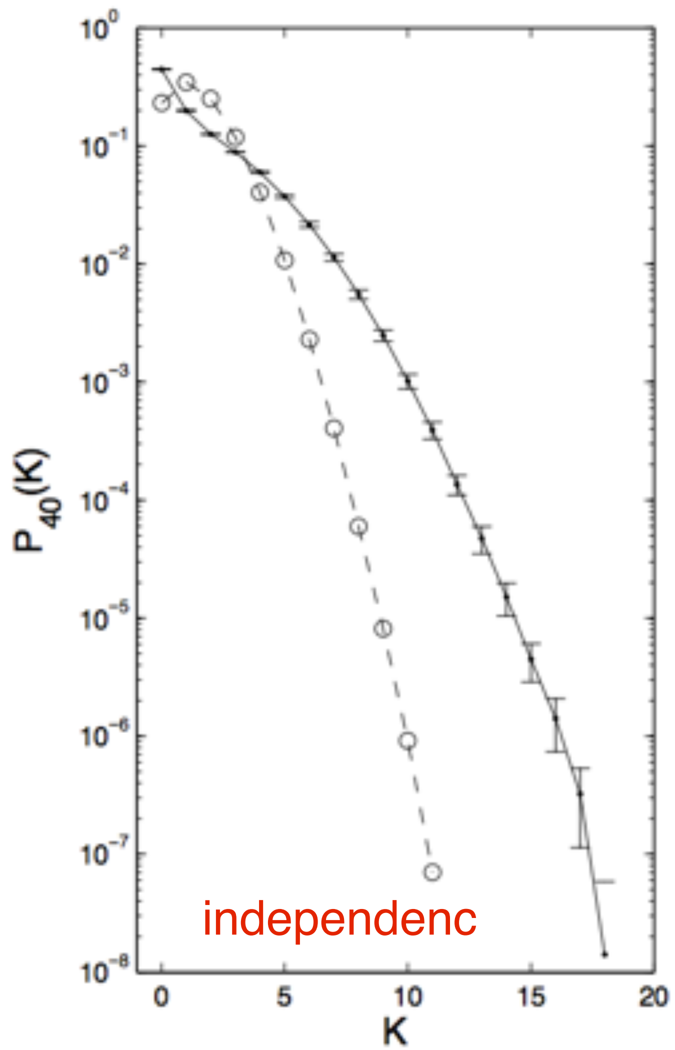
- raster → binary variables $\sigma_i = 0, 1$ N ~ 150 neurons





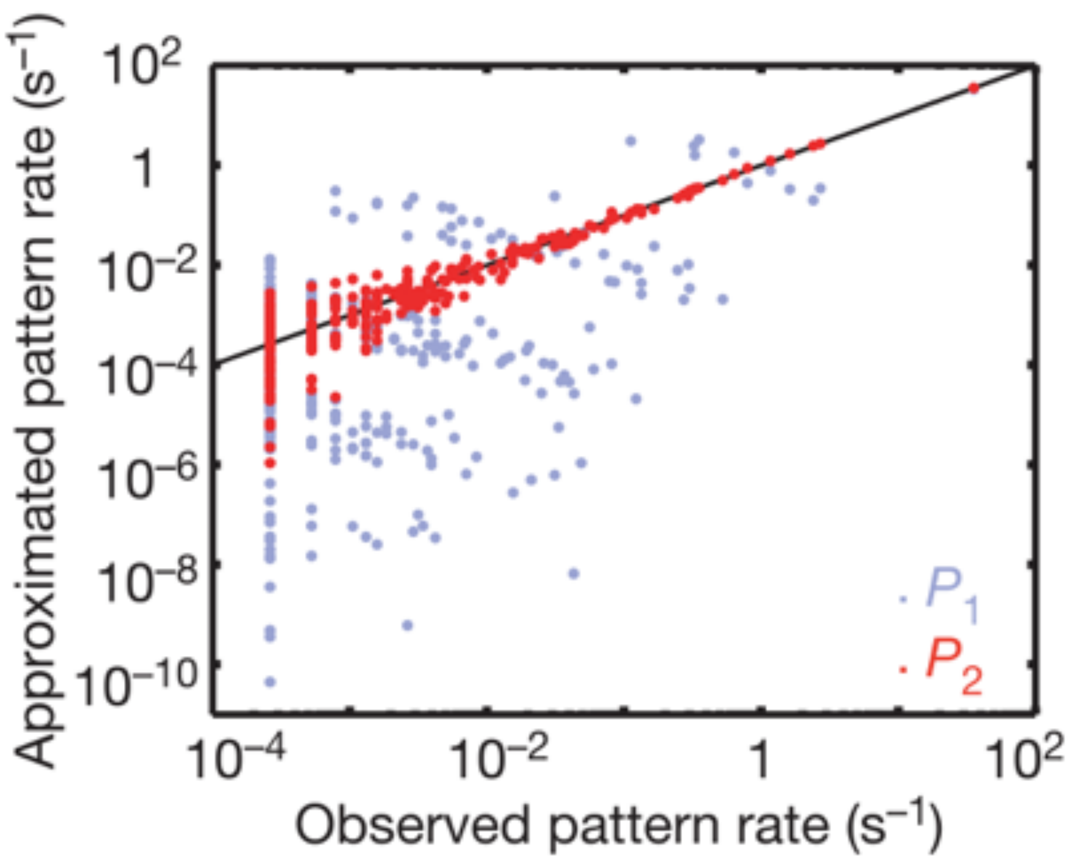
Marre et al.
J. Neurosci. 2012

neuron activities are correlated



total number of spikes

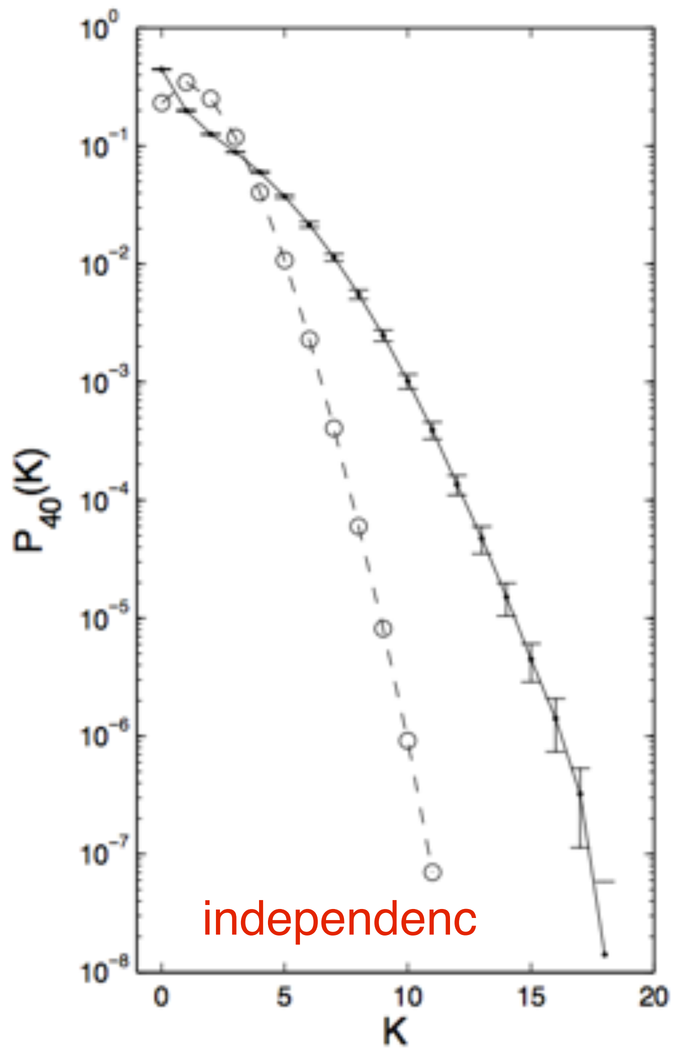
Schneidman et al, Nature 2005



$$P_1(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_i h_i \sigma_i}$$

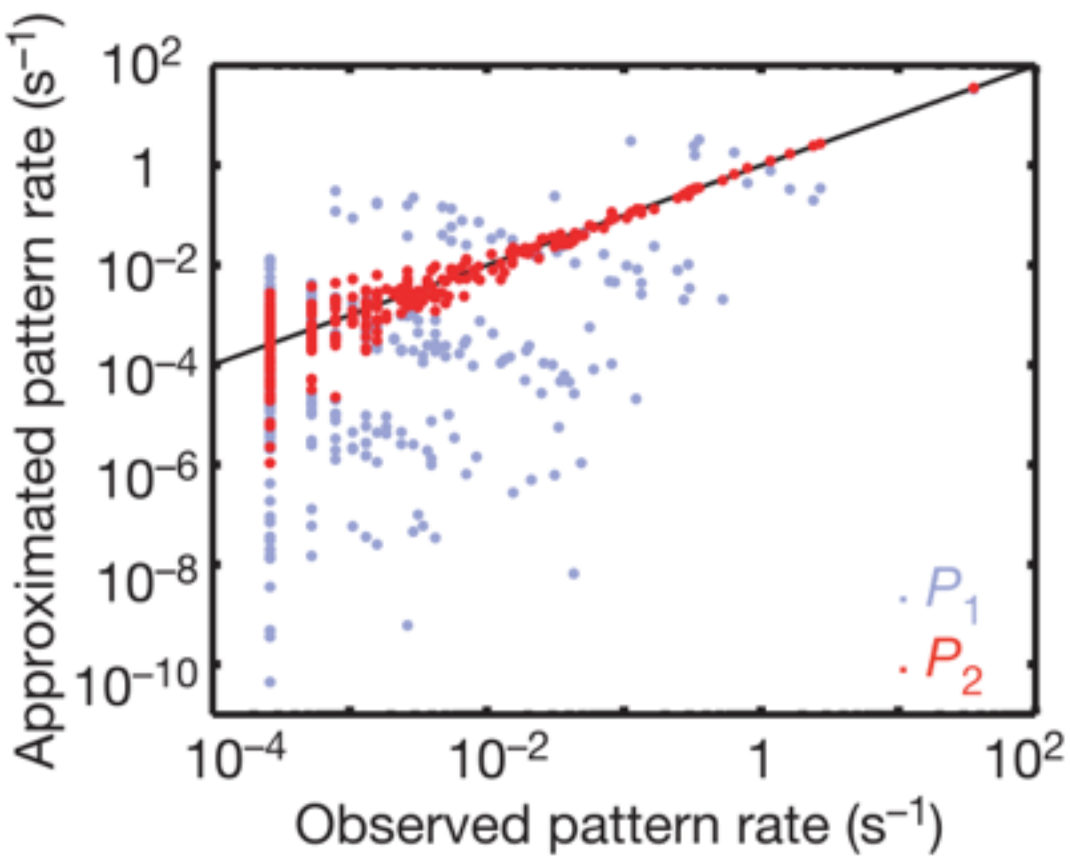
Ising model $$P_2(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_i h_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j}$$

neuron activities are correlated



total number of spikes

Schneidman et al, Nature 2005



$$P_1(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_i h_i \sigma_i}$$

Ising model $$P_2(\boldsymbol{\sigma}) = \frac{1}{Z} e^{\sum_i h_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j}$$

goal: build the thermodynamics of this correlated system from data

building the density of states

- evaluate $P(\sigma_1, \ldots, \sigma_N)$ by modelling or by frequency counting

building the density of states

- evaluate $P(\sigma_1, \ldots, \sigma_N)$ by modelling or by frequency counting
- define "energy" through Boltzmann law

$$P = \frac{1}{Z} e^{-E/k_B T} \longrightarrow E = -\log P$$

building the density of states

- evaluate $P(\sigma_1, \ldots, \sigma_N)$ by modelling or by frequency counting
- define "energy" through Boltzmann law

$$P = \frac{1}{Z} e^{-E/k_B T} \longrightarrow E = -\log P$$

- now consider the *distribution* of energies *E*

$$C(E) = \text{number of states with } E(\sigma) < E$$

building the density of states

- evaluate $P(\sigma_1, \ldots, \sigma_N)$ by modelling or by frequency counting
- define "energy" through Boltzmann law

$$P = \frac{1}{Z} e^{-E/k_B T} \longrightarrow E = -\log P$$

- now consider the *distribution* of energies *E*

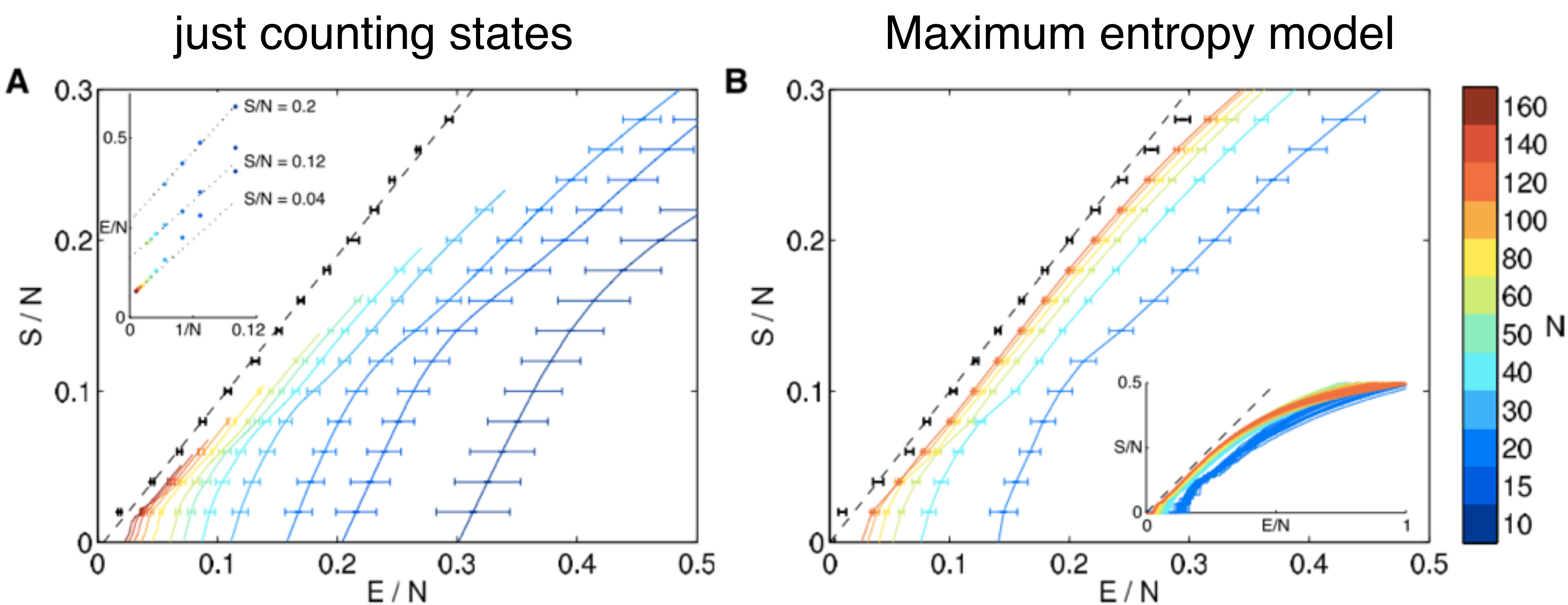
$$C(E) = \text{number of states with } E(\sigma) < E$$

- define a microcanonical entropy :

$$S(E) = \log C(E)$$

Mora Bialek J. Stat. Phys. 2011

density of states



(under natural movie stimulus)

Tkacik Mora Marre Amodei Berry Bialek, arxiv 2014

Zipf's law (interlude)

HUMAN BEHAVIOR

AND

THE PRINCIPLE OF LEAST EFFORT

An Introduction to Human Ecology

by

GEORGE KINGSLEY ZIPF, Ph.D.
Harvard University

Zipf 1949

Zipf's law (interlude)

Probability P (*E* in log scale)

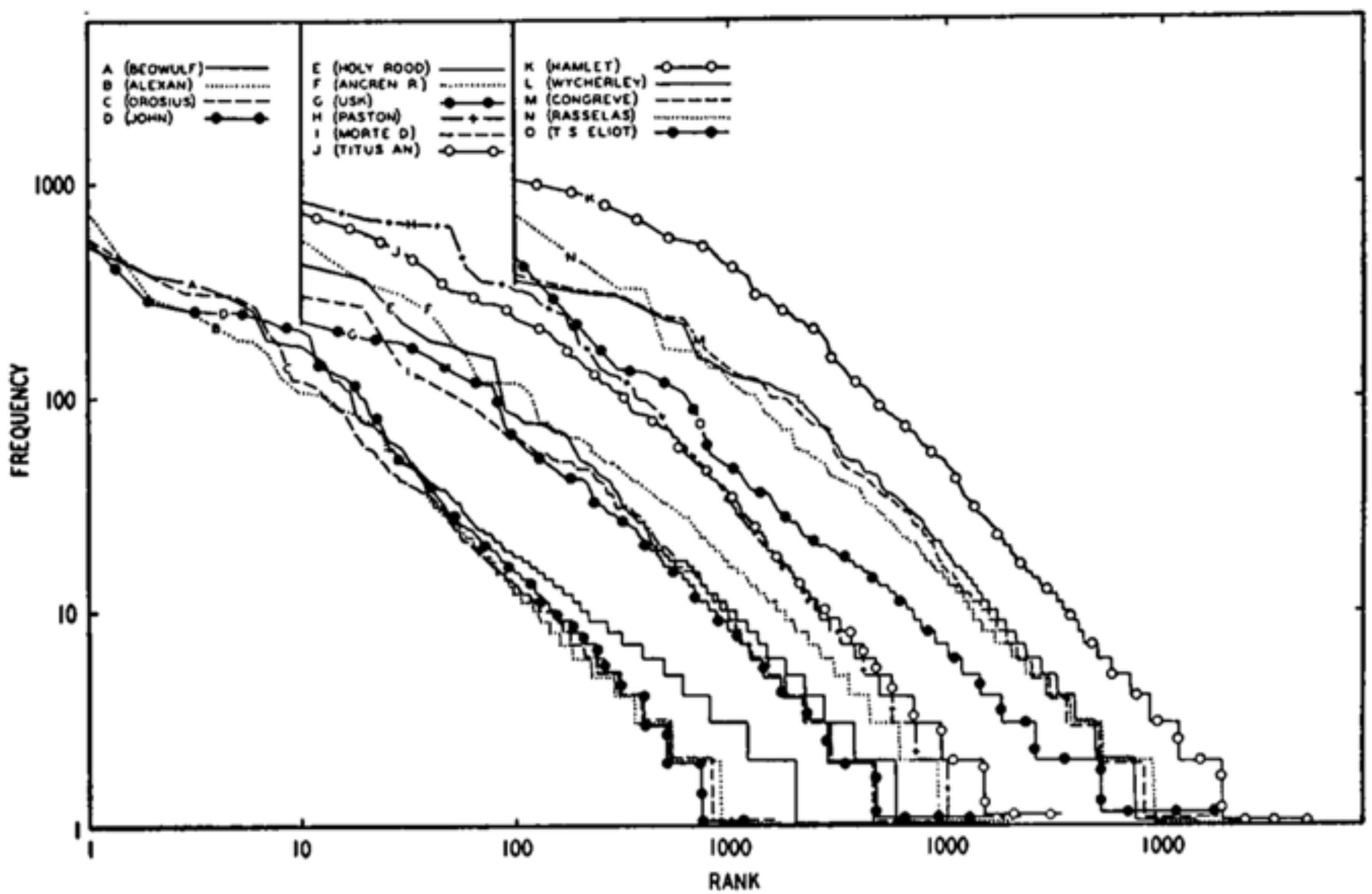


Fig. 3–14. Beowulf to T. S. Eliot. Rank-frequency distributions of the words of fifteen English writers from early Old English to the present day.

~ cumulative distribution

Zipf 1949

what does $S(E) = E$ mean?

what does $S(E) = E$ mean?

what's the probability of a given energy?

how many states at *E* → ← probability of a state at *E*

$$P(E) \approx e^{S-E}$$

constant!

what does $S(E) = E$ mean?

what's the probability of a given energy?

how many states at E → probability of a state at E

$$P(E) \approx e^{S-E}$$

constant!

(note: only works if exponent is = **1**)

what does $S(E) = E$ mean?

what's the probability of a given energy?

how many states at E → ← probability of a state at E

$$P(E) \approx e^{S-E}$$

constant!

(note: only works if exponent is = 1)

in usual thermodynamics…

- E scales with N
- its fluctuations scale with N

heat capacity $C = \mathrm{Var}(E) \sim N$

C / N diverges at 2nd order transition critical point (e.g. 2D Ising model)

what does *S(E)* = *E* mean?

what's the probability of a given energy?

how many states at *E* → probability of a state at *E*

$$P(E) \approx e^{S-E}$$

constant!

(note: only works if exponent is = 1)

in usual thermodynamics…

- E scales with N
- its fluctuations scale with N

heat capacity $C = \mathrm{Var}(E) \sim N$

C / N diverges at 2nd order transition critical point (e.g. 2D Ising model)

link to information theory $E = -\log P$ "surprise" (Shannon 1948)

- equipartition theorem (valid for independent units): almost all codewords we see have the same surprise ~ entropy
- basis for compression

specific heat

let's add a spurious temperature — one direction in parameter space

$$P_T(\sigma) = \frac{1}{Z(T)} e^{-E/T} \qquad C = \mathrm{Var}_T(E/T) = \mathrm{Var}_T(-\log P)$$

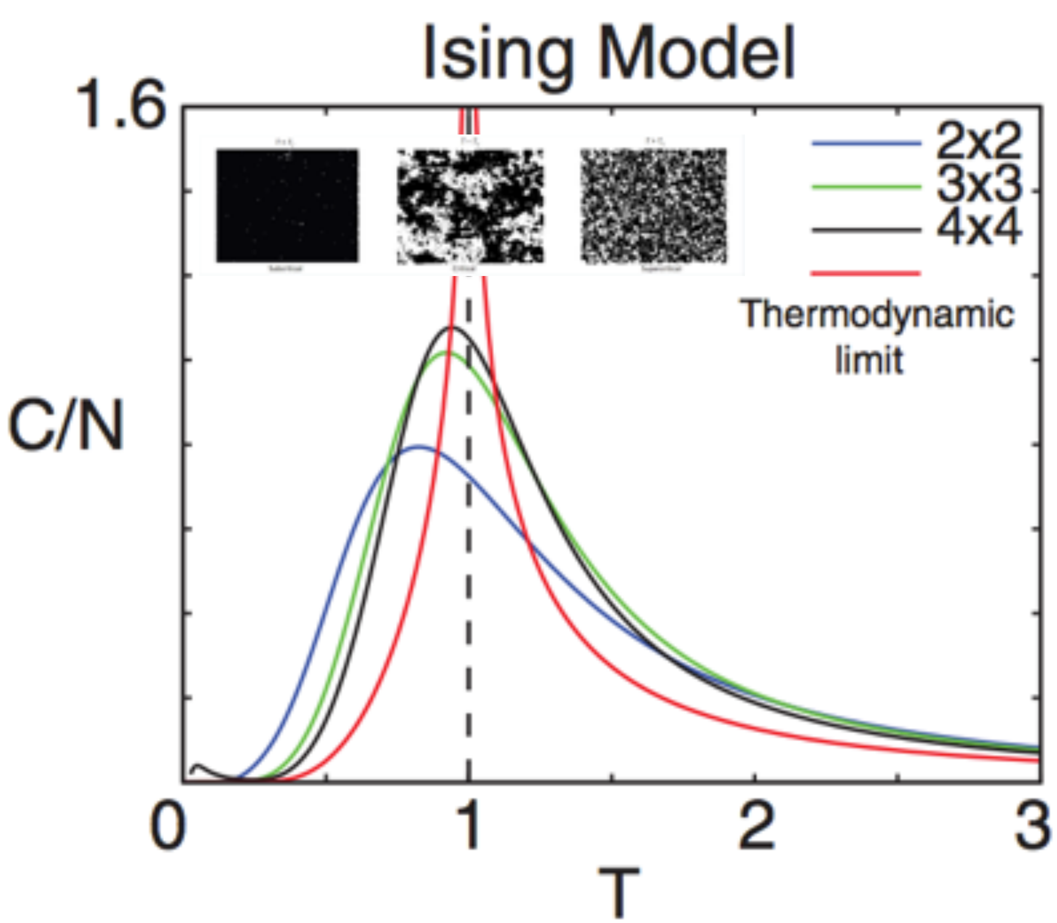
(T = 1 corresponds to the real ensemble)

specific heat

let's add a spurious temperature — one direction in parameter space

$$P_T(\sigma) = \frac{1}{Z(T)} e^{-E/T} \qquad C = \mathrm{Var}_T(E/T) = \mathrm{Var}_T(-\log P)$$

($T = 1$ corresponds to the real ensemble)

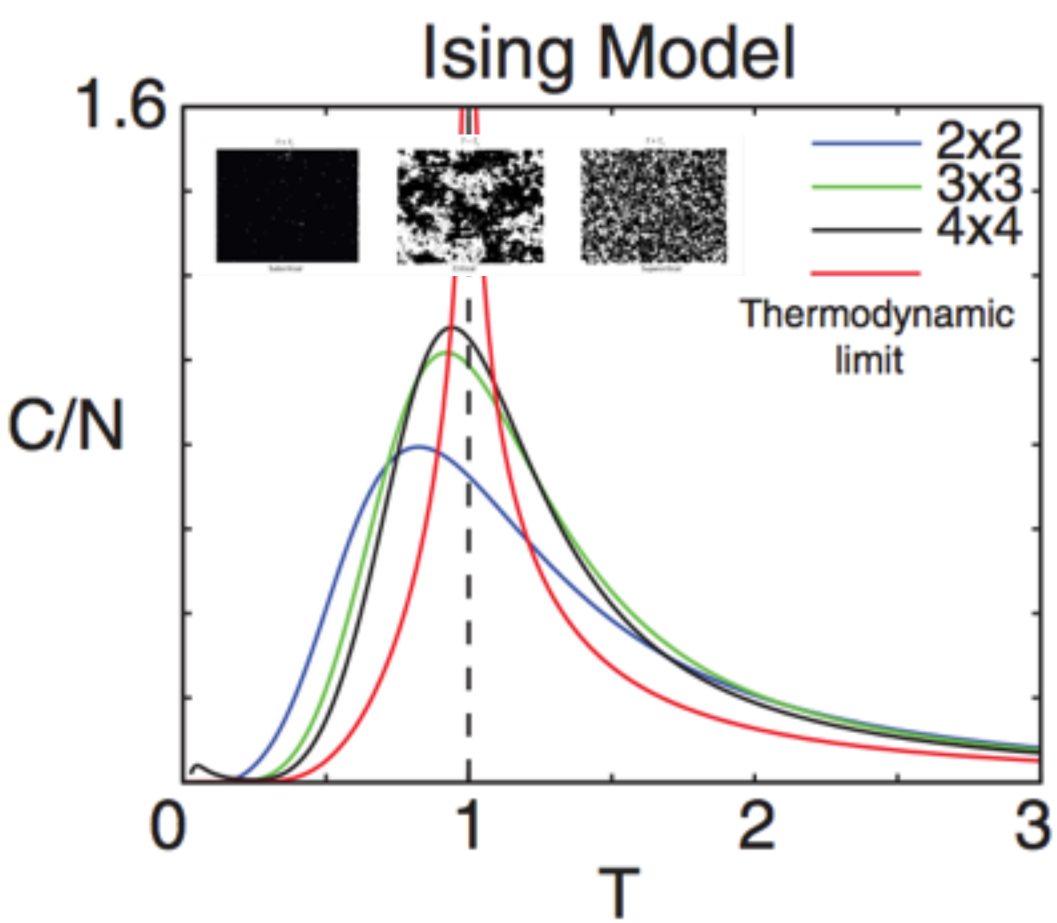


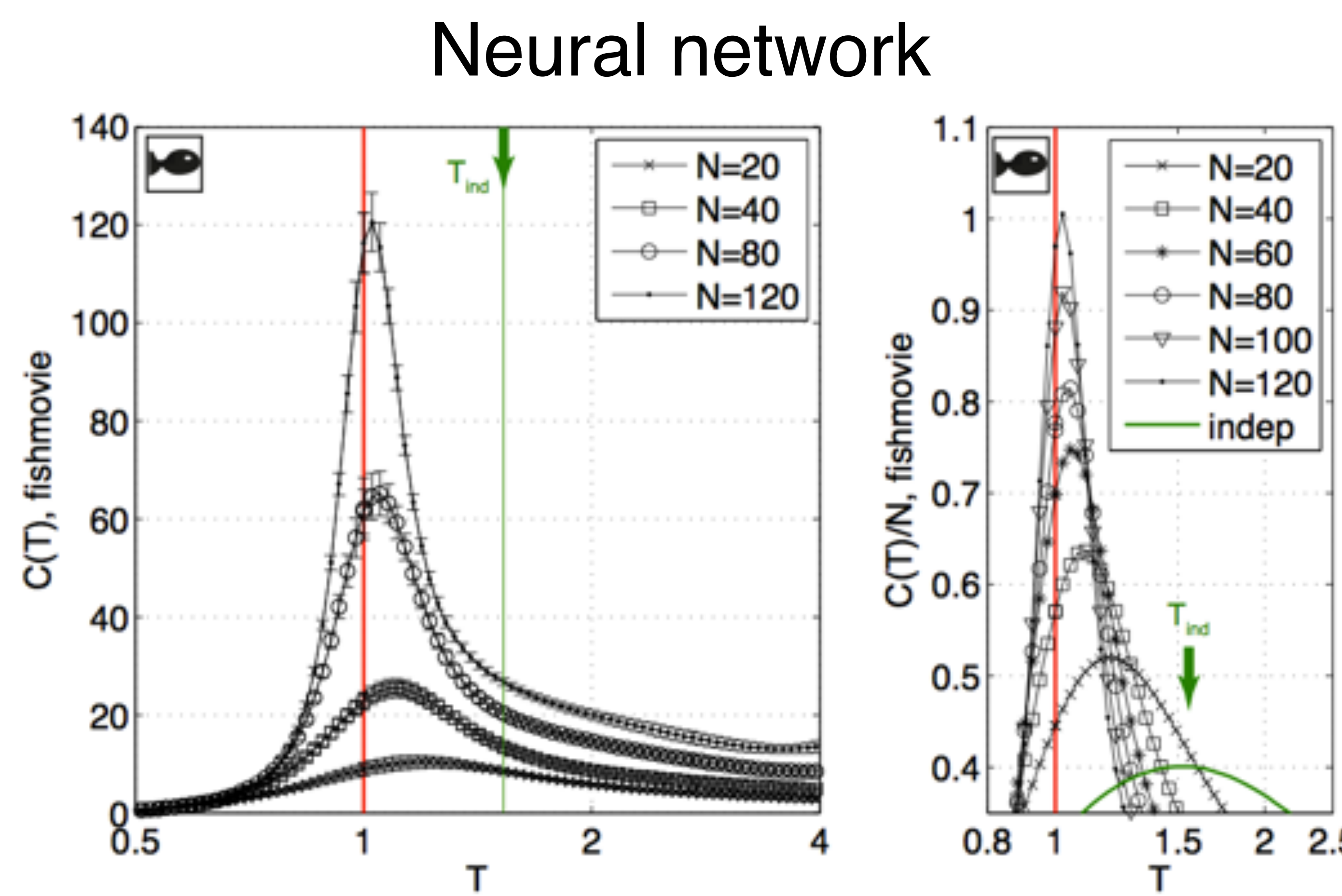
specific heat

let's add a spurious temperature — one direction in parameter space

$$P_T(\sigma) = \frac{1}{Z(T)} e^{-E/T} \qquad C = \mathrm{Var}_T(E/T) = \mathrm{Var}_T(-\log P)$$

(T = 1 corresponds to the real ensemble)





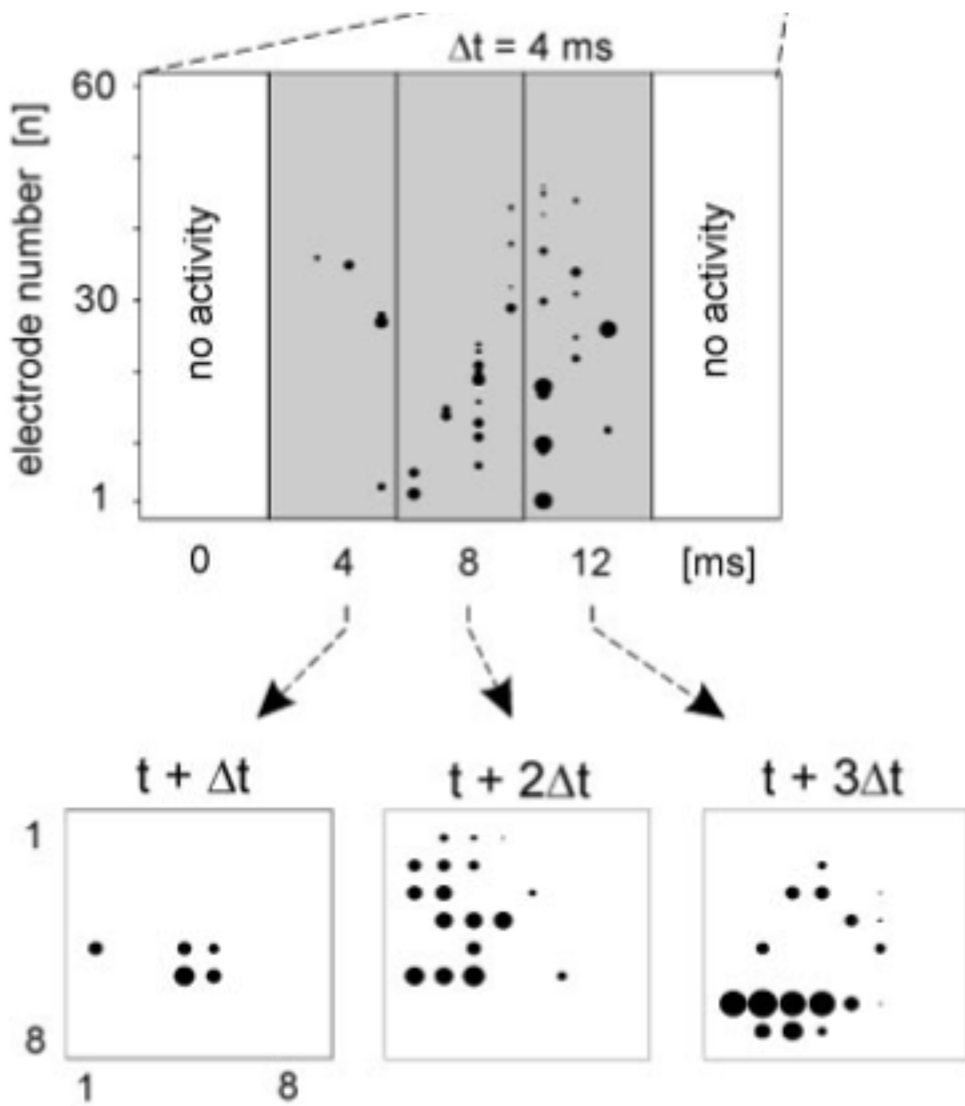
Tkacik Mora Marre Amodei Berry Bialek, arxiv 2014

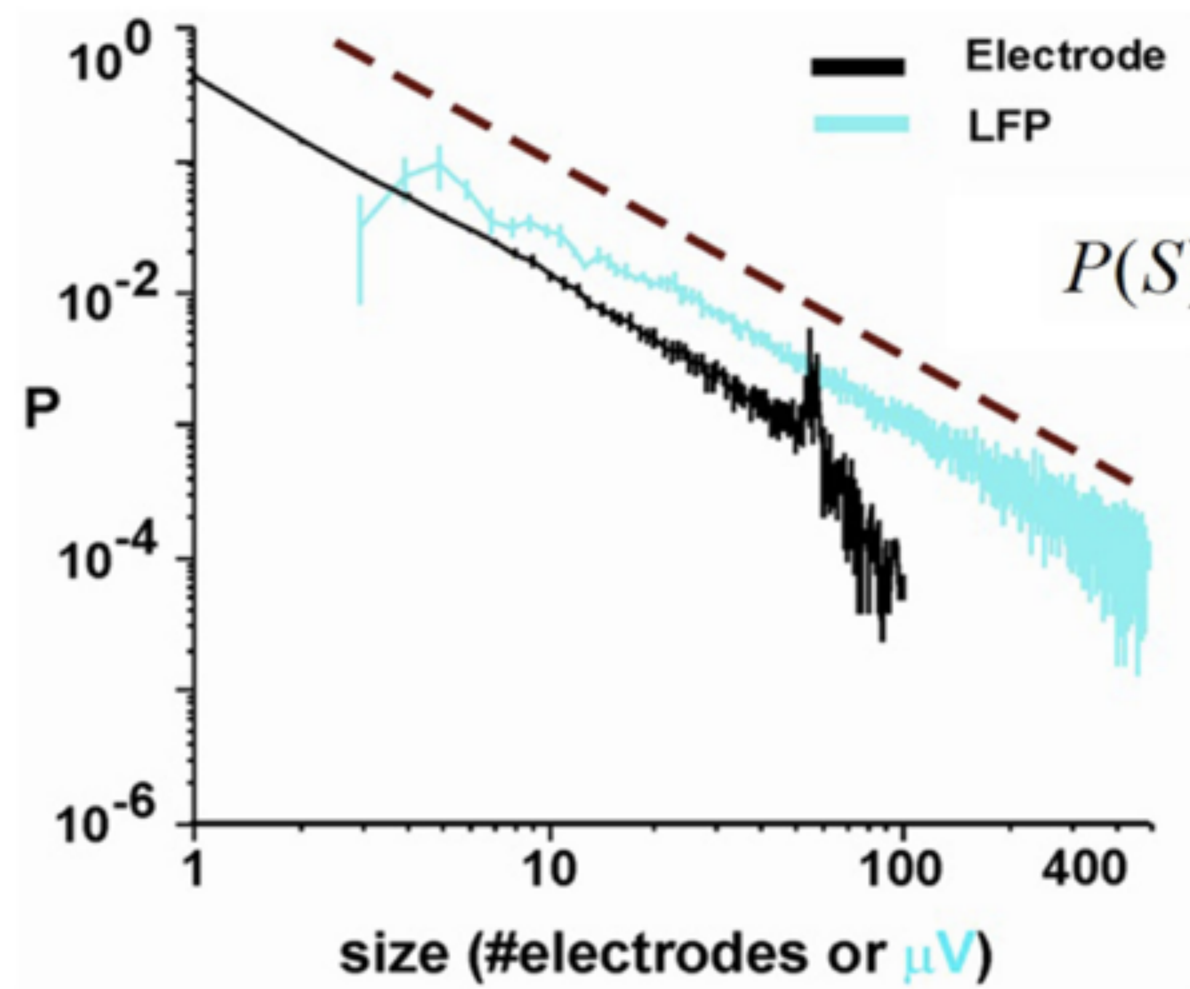
dynamical criticality

The Journal of Neuroscience, December 3, 2003 • 23(35):11167–11177 • **11167**

Neuronal Avalanches in Neocortical Circuits

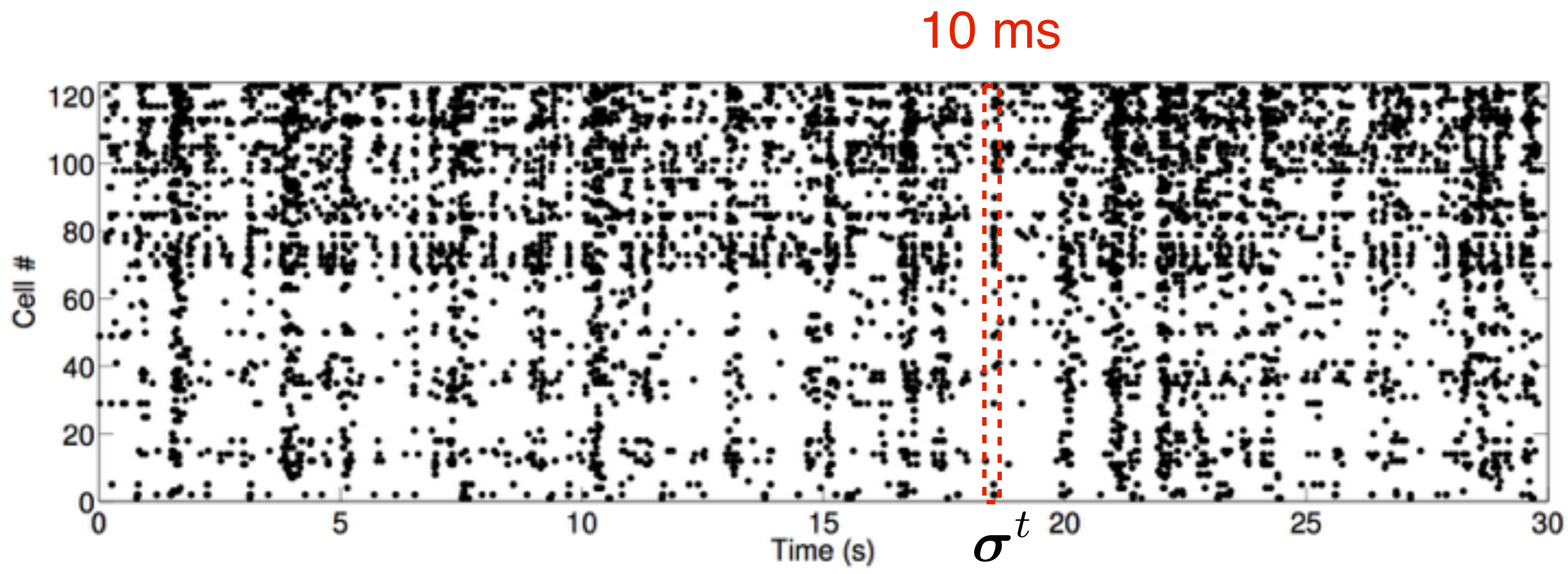
John M. Beggs and Dietmar Plenz
Unit of Neural Network Physiology, Laboratory of Systems Neuroscience, National Institute of Mental Health, Bethesda, Maryland 20892



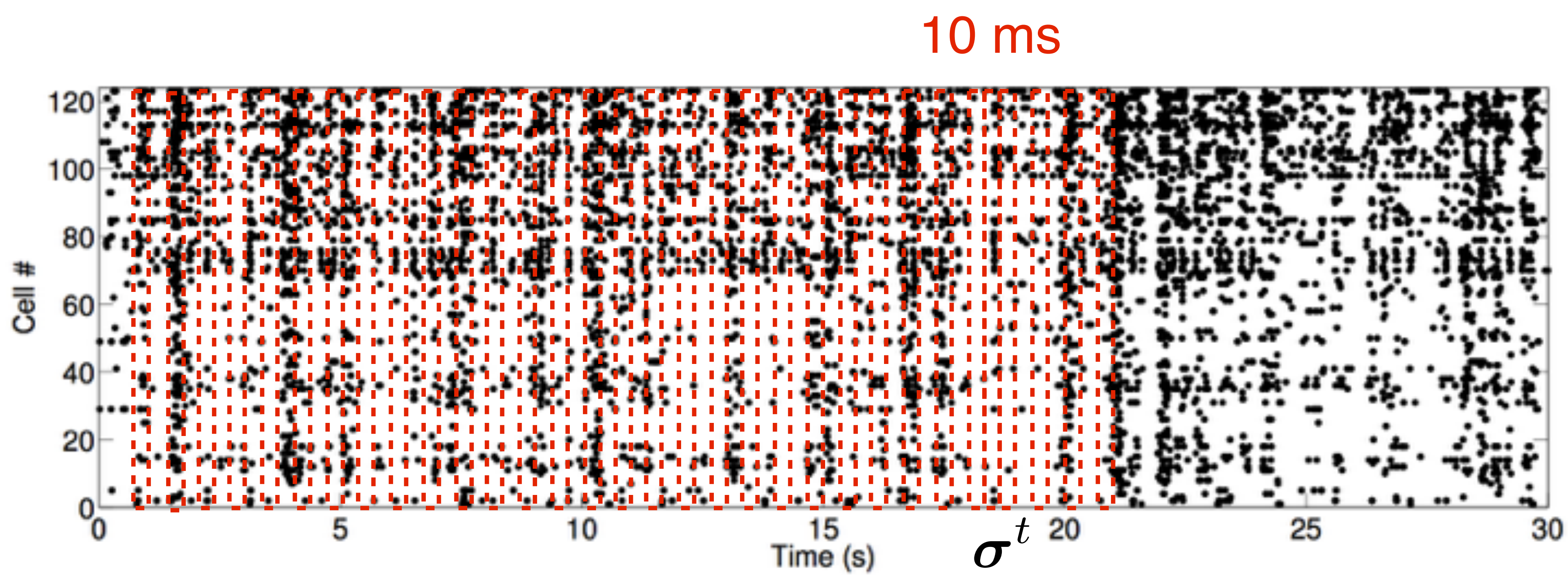


$$P(S) \sim S^{-\alpha}$$

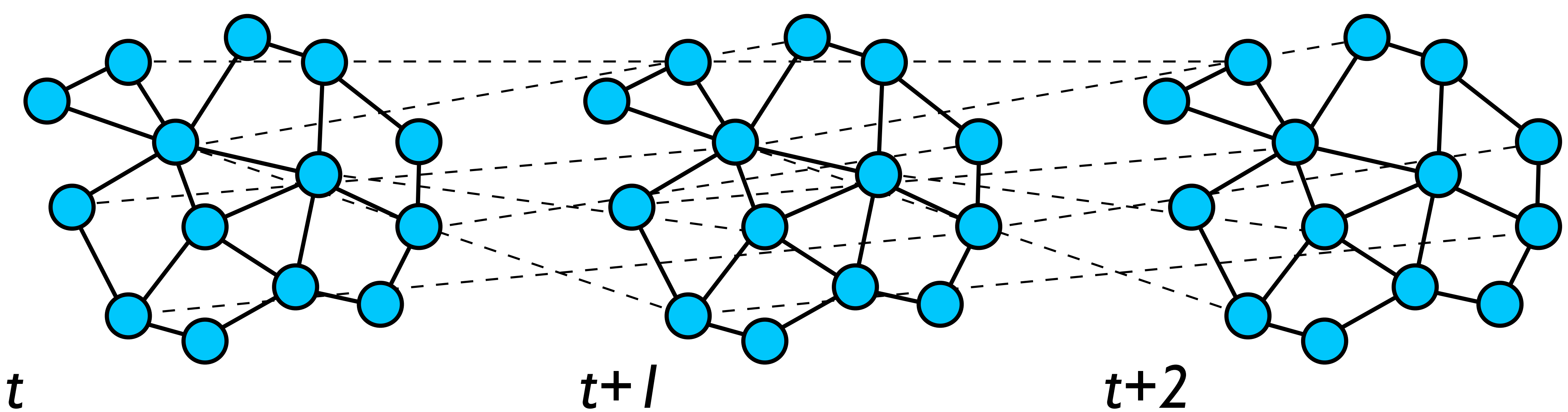
dynamical approach



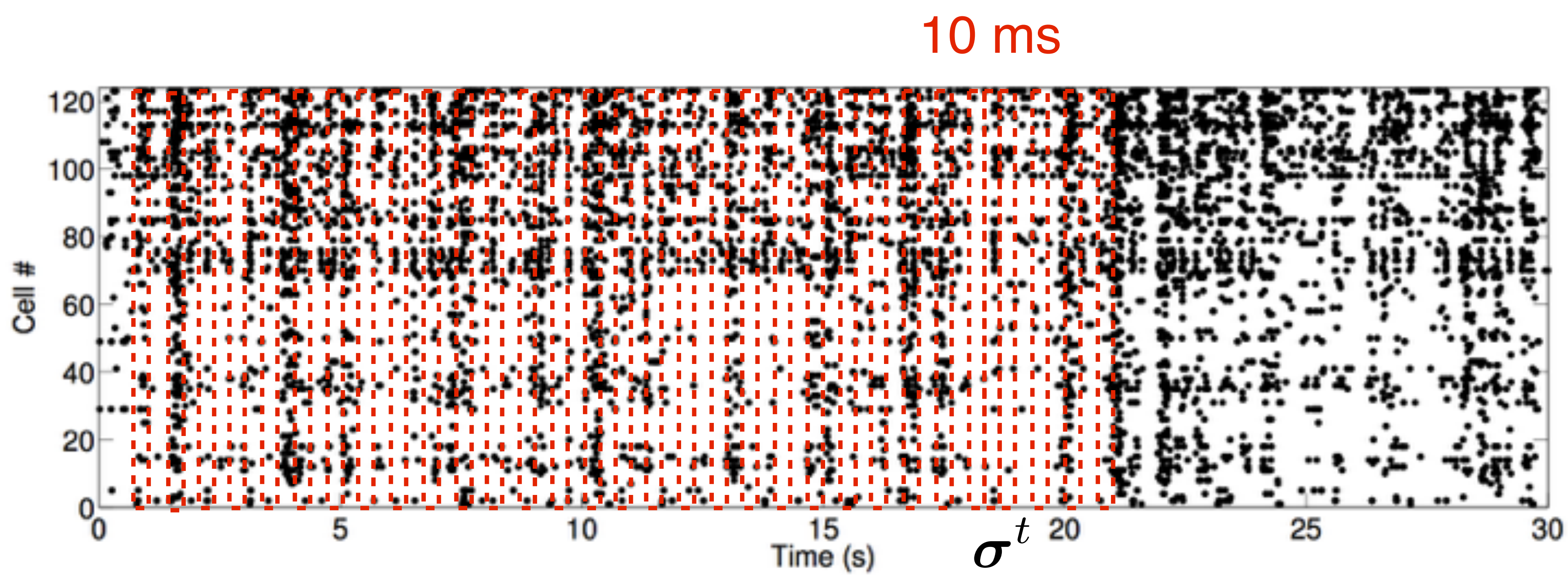
dynamical approach



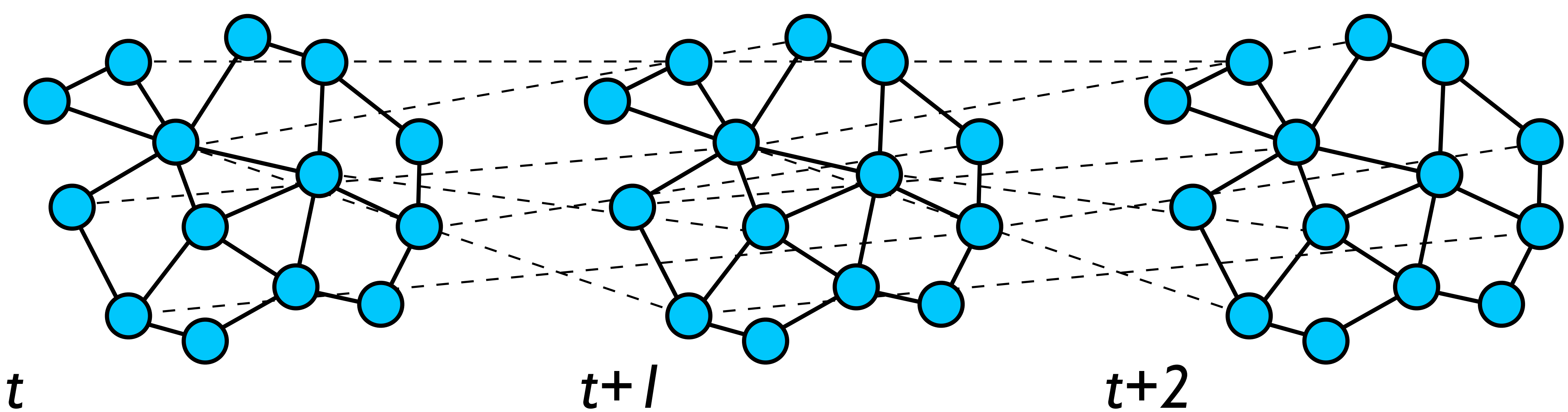
- proposal: consider statistics over trajectories $P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^L)$



dynamical approach

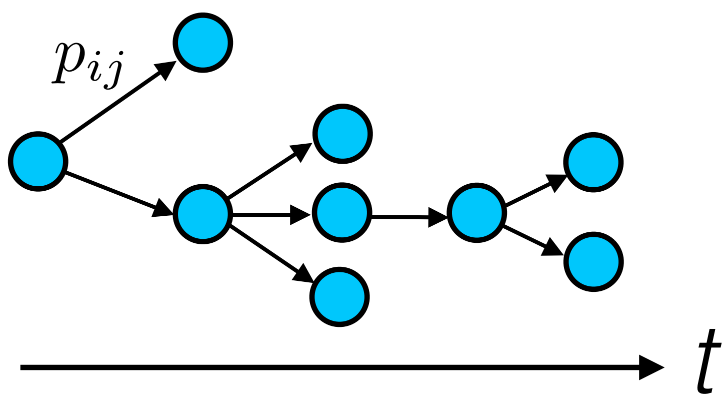


- proposal: consider statistics over trajectories $P(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^L)$



- define $E = -\log P(\{\sigma_{i,t}\})$
- calculate specific heat $c = \frac{1}{NL}\mathrm{Var}(E)$

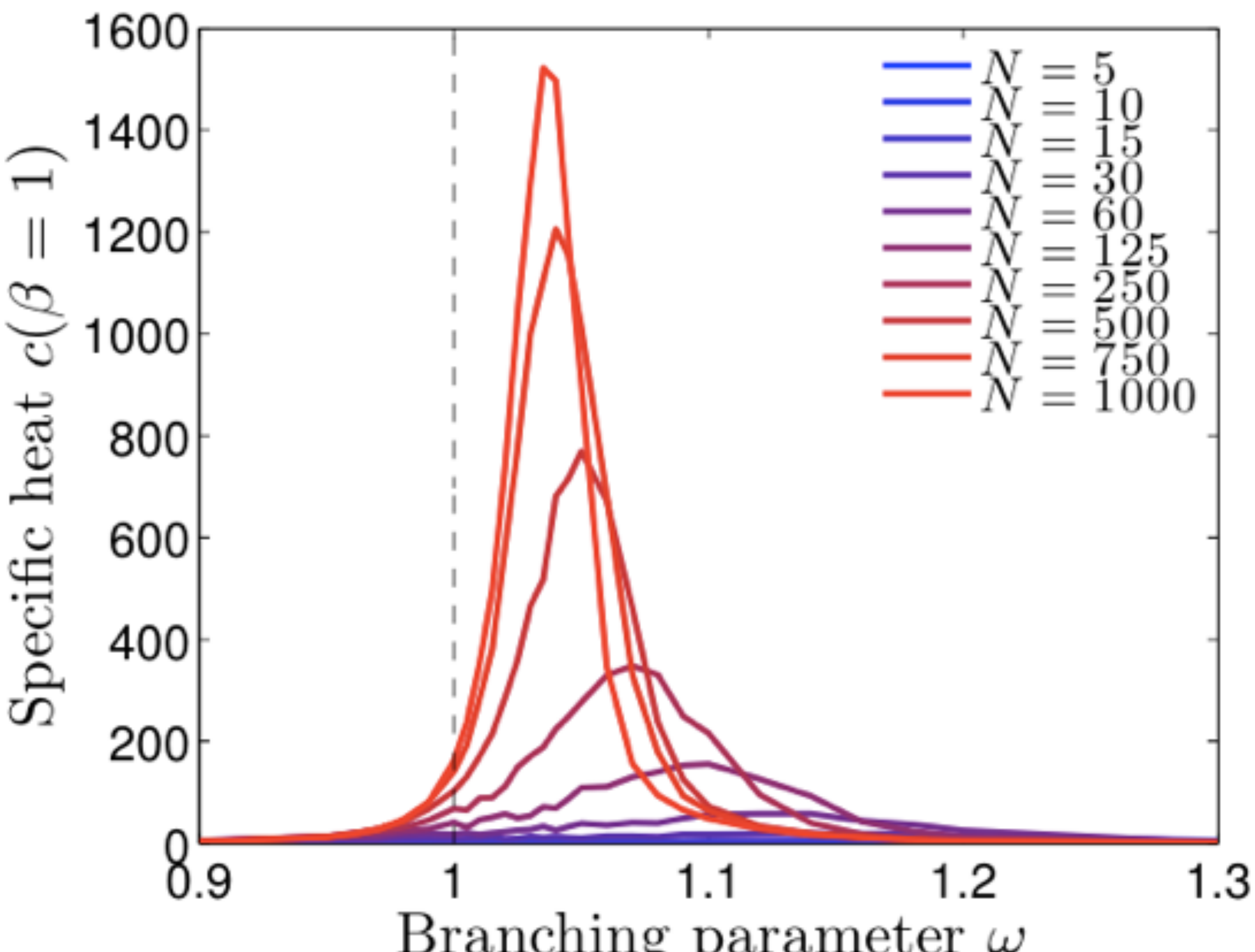
link to dynamical criticality: branching process



$$P(\{\sigma_{i,t}\}) = \prod_t \prod_{i=1}^{N} p_i(t)^{\sigma_{i,t}} [1 - p_i(t)]^{1-\sigma_{i,t}}$$

$$p_i(t) = 1 - \prod_j (1 - p_{ij})^{\sigma_{i,t-1}}$$

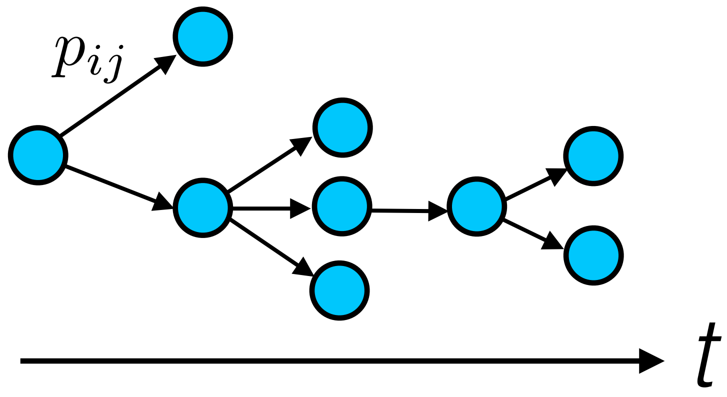
Beggs & Plenz 2003



branching parameter:

$$\omega = \frac{1}{N} \sum_{ij} p_{ij}$$

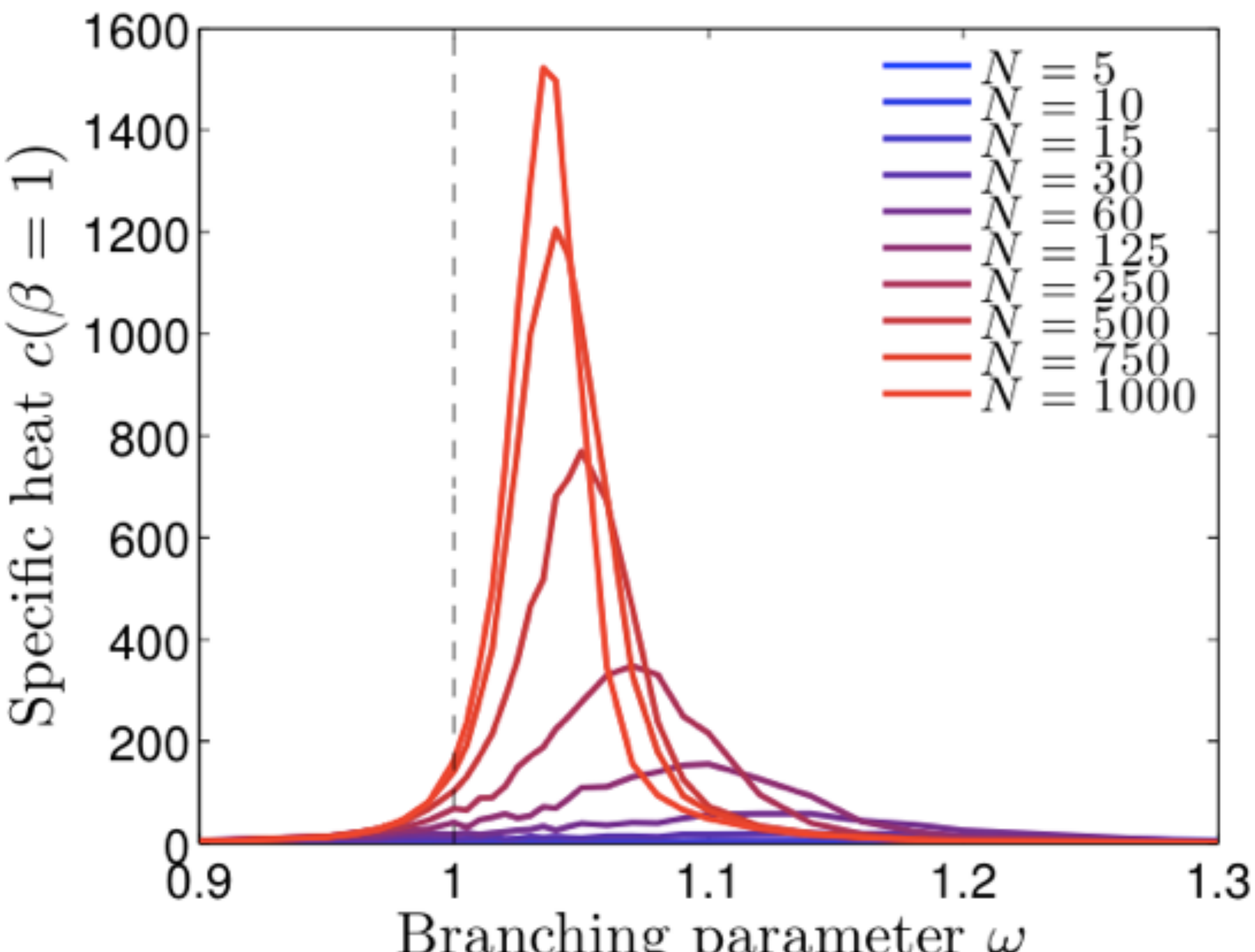
link to dynamical criticality: branching process



$$P(\{\sigma_{i,t}\}) = \prod_t \prod_{i=1}^{N} p_i(t)^{\sigma_{i,t}} [1 - p_i(t)]^{1-\sigma_{i,t}}$$

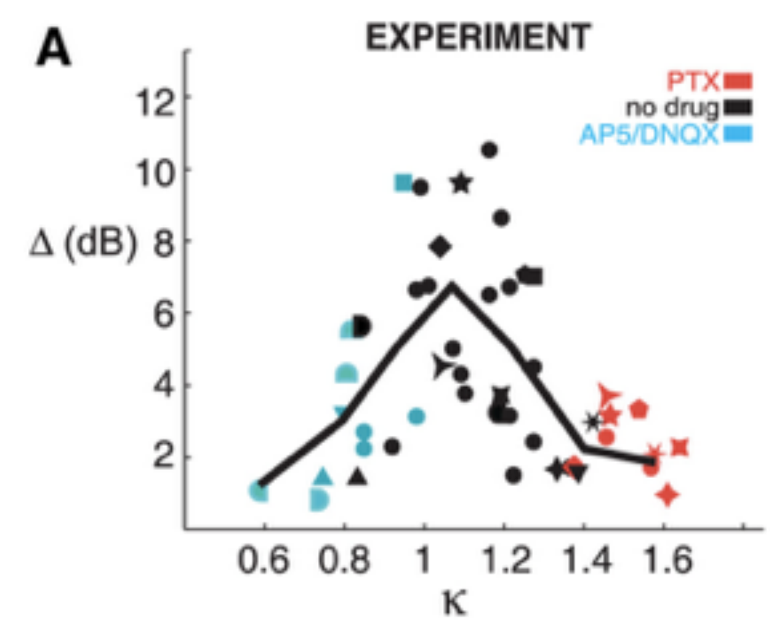
$$p_i(t) = 1 - \prod_j (1 - p_{ij})^{\sigma_{i,t-1}}$$

Beggs & Plenz 2003



branching parameter:

$$\omega = \frac{1}{N} \sum_{ij} p_{ij}$$



Shew Yang Petermann Roy Plenz 2009

model for multi-neuron spike trains

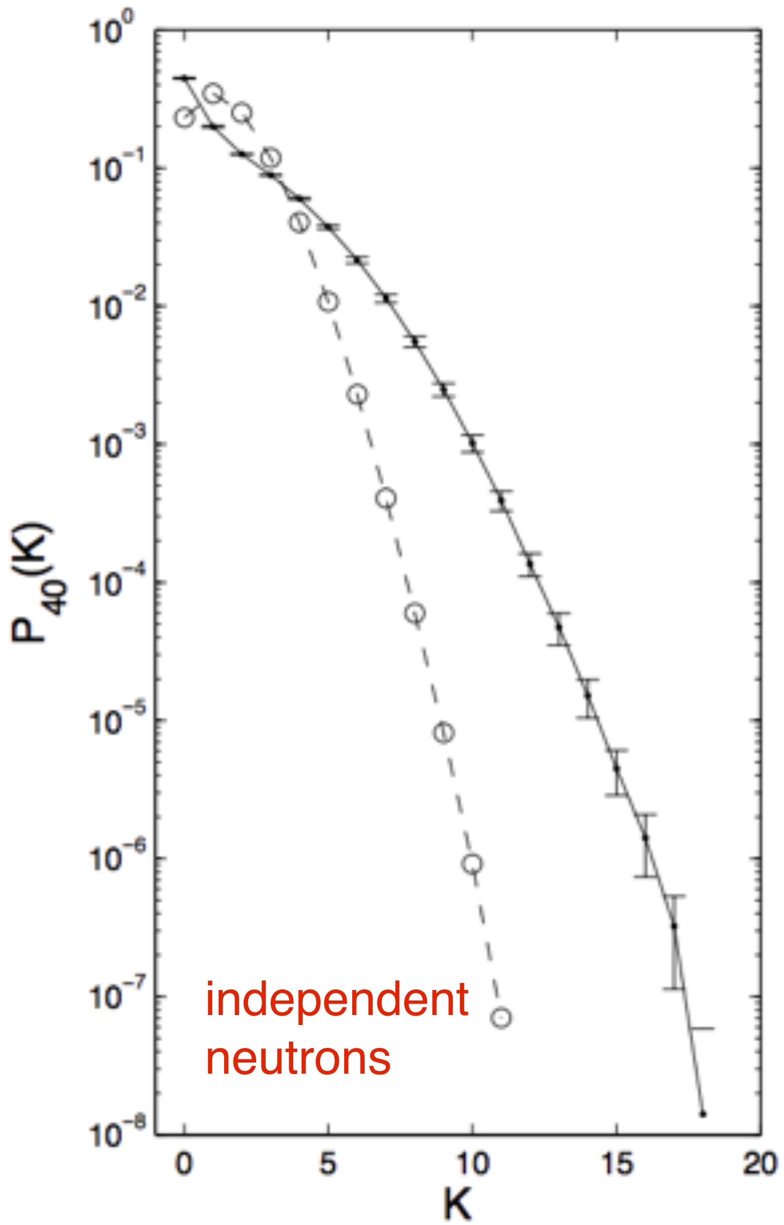
sampling 2^N states is hard enough; here 2^{NL} states — we need models

let's do something simple

- total number of spikes $$K_t = \sum_i \sigma_{i,t}$$

is informative of collective behaviour

Tkacik Marre Mora Amodei Berry Bialek, JSTAT 2013



model for multi-neuron spike trains

sampling 2^N states is hard enough; here 2^{NL} states — we need models

let's do something simple

- total number of spikes $K_t = \sum_i \sigma_{i,t}$

is informative of collective behaviour

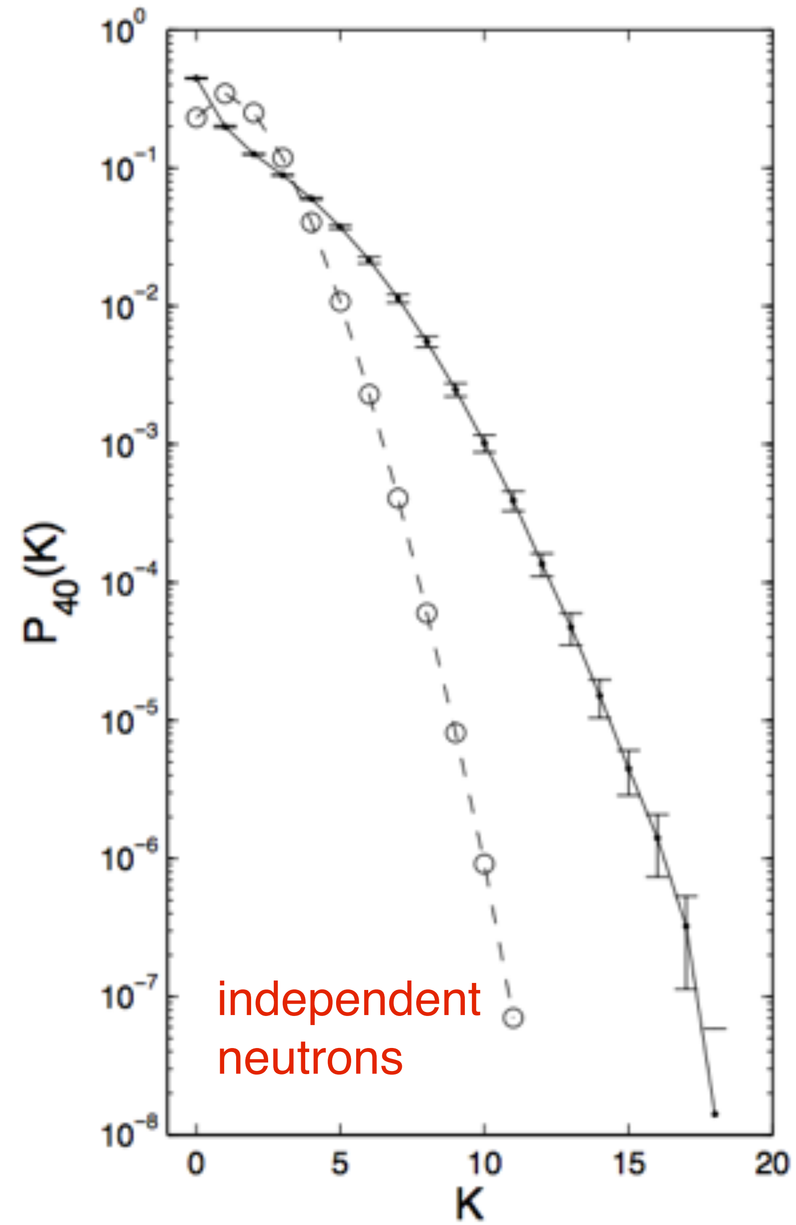
Tkacik Marre Mora Amodei Berry Bialek, JSTAT 2013

- maximum entropy model with

constrains on temporal correlations of *K*

$$P(K_t, K_{t'}) \qquad |t - t'| < v$$

⇔ all neurons behave the same



model for multi-neuron spike trains

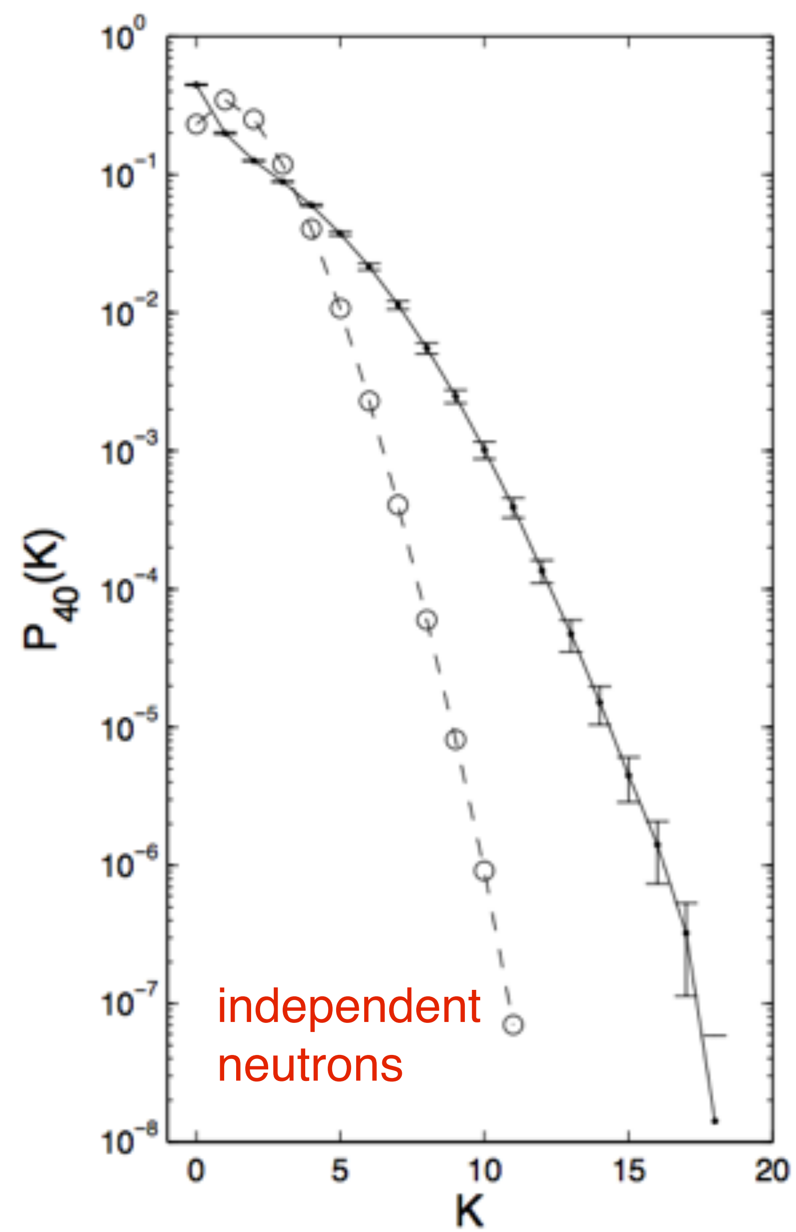
sampling 2^N states is hard enough; here 2^{NL} states — we need models

let's do something simple

- total number of spikes $K_t = \sum_i \sigma_{i,t}$

is informative of collective behaviour

Tkacik Marre Mora Amodei Berry Bialek, JSTAT 2013



- maximum entropy model with

constrains on temporal correlations of *K*

$$P(K_t, K_{t'}) \qquad |t - t'| < v$$

⇔ all neurons behave the same

- "Energy"

$$E = -\sum_t h(K_t) - \sum_t \sum_{u=1}^{v} J_u(K_t, K_{t+u})$$

solving the problem

$$E = -\sum_t h(K_t) - \sum_t \sum_{u=1}^{v} J_u(K_t, K_{t+u})$$

- define a "super-variable"

$$X_t = (K_t, K_{t+1}, \ldots, K_{t+v-1})$$

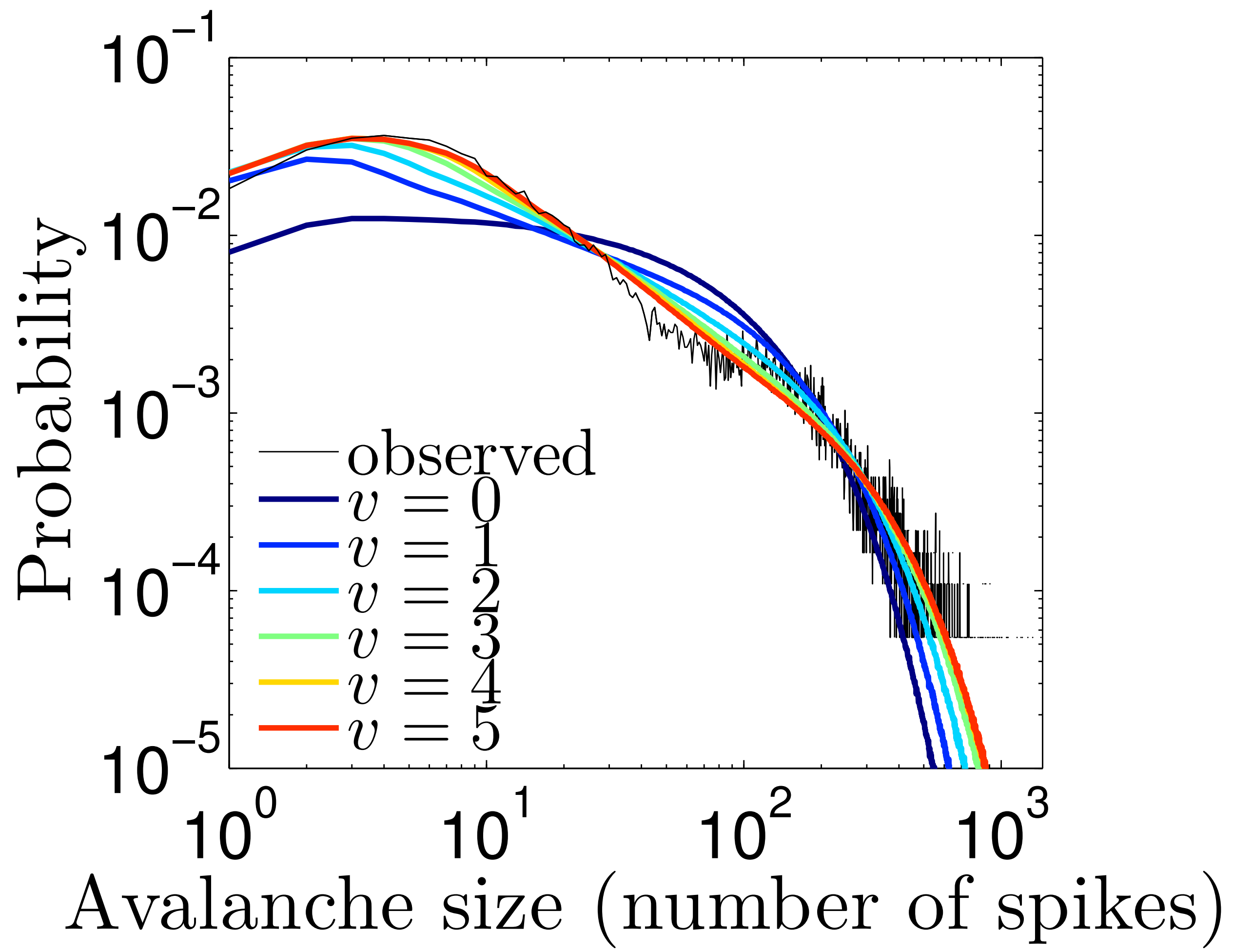
- now becomes a 1D model

$$P(\{X_t\}) = \frac{1}{Z} \exp\left[\sum_t H(X_t) + \sum_t W(X_t, X_{t+1})\right]$$

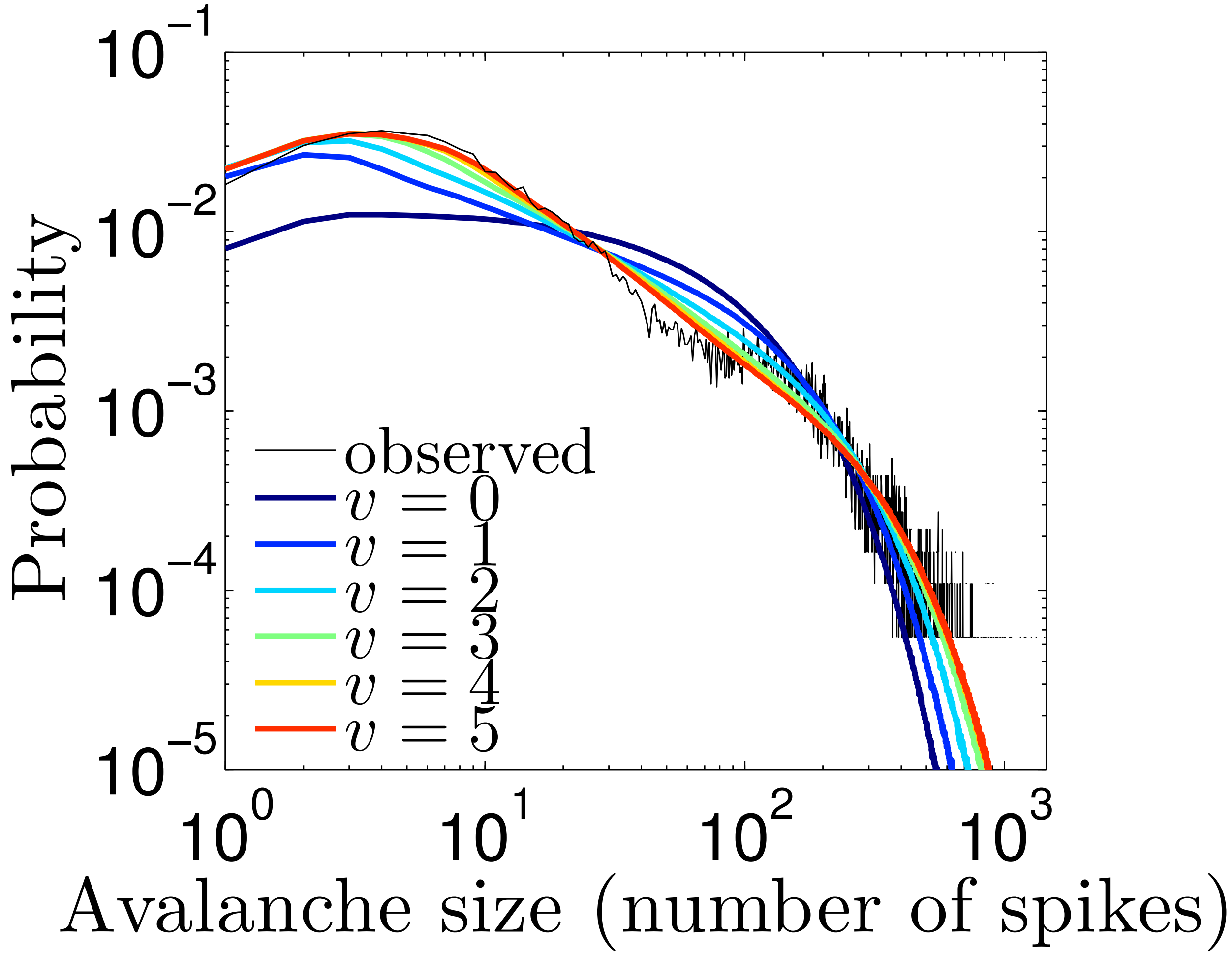
- can be solved by transfer matrices

  (*aka* forward backward algorithm, or belief propagation in 1D)

model predicts avalanche dynamics

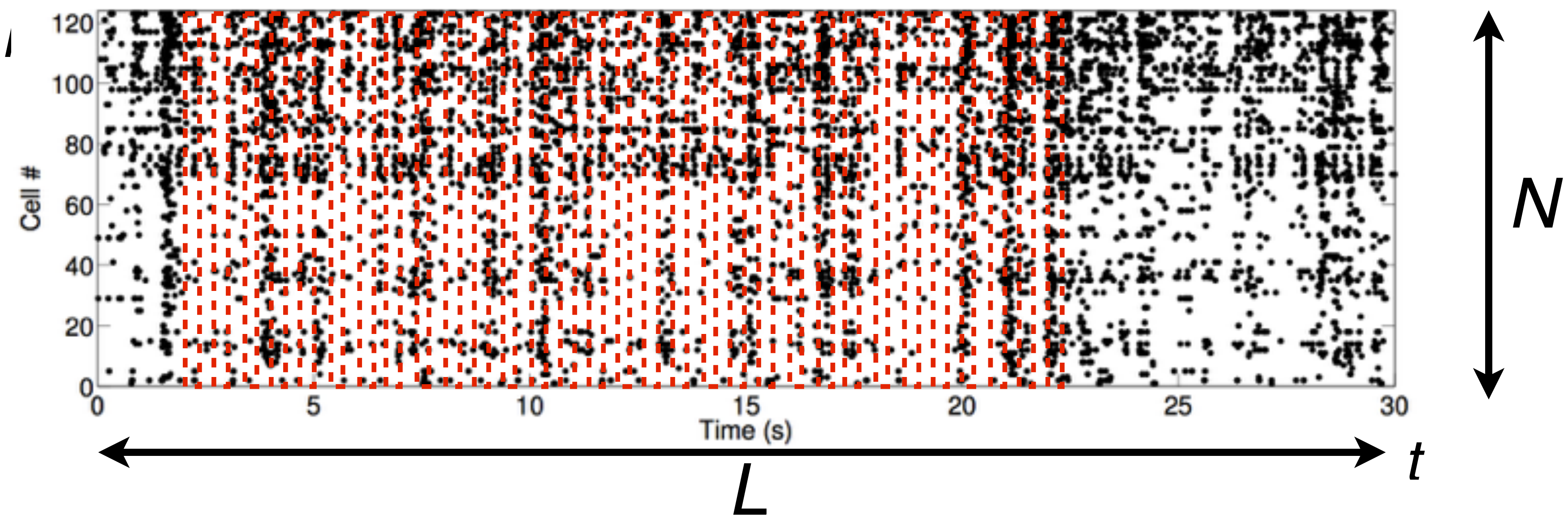


model predicts avalanche dynamics



NB: no power laws in avalanche statistics

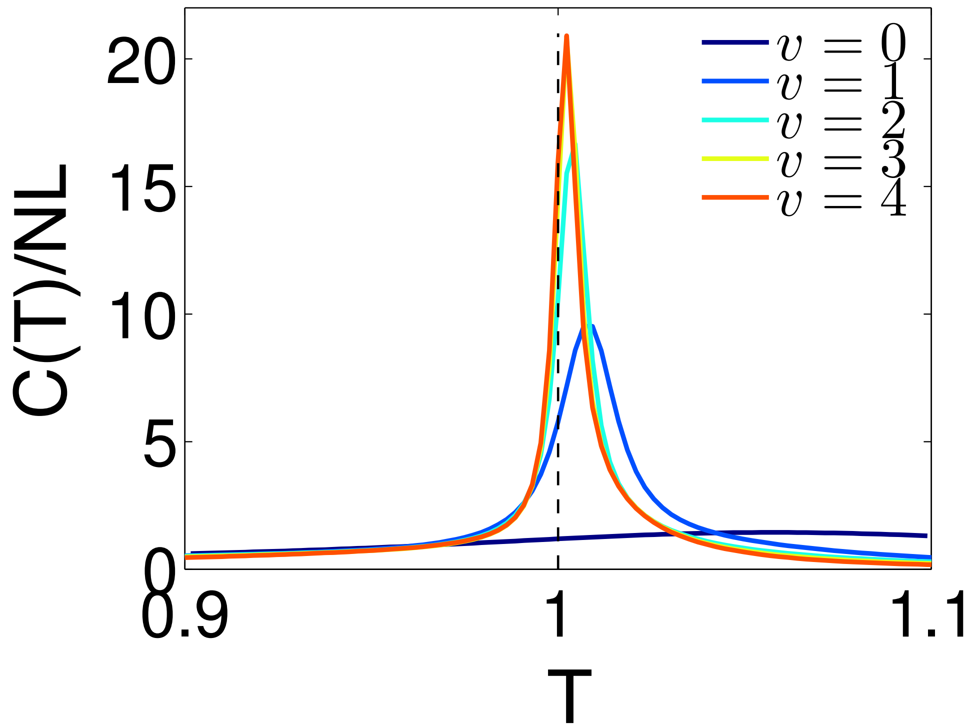
thermodynamics of spike trains

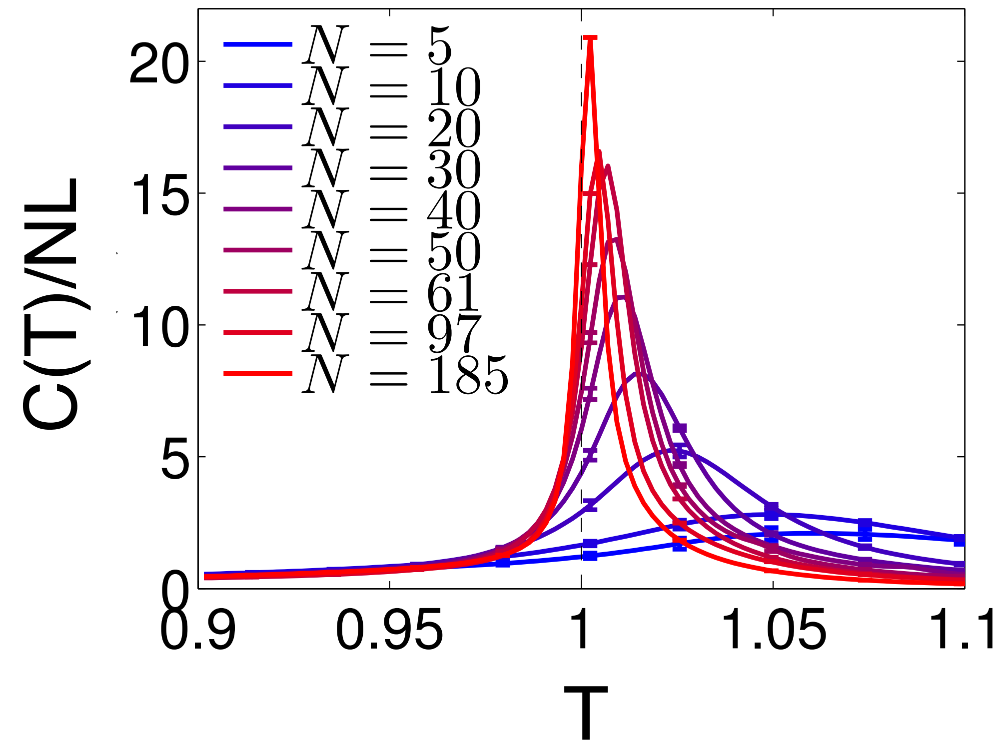


$$E = -\log P(\{\sigma_{i,t}\})$$

$$P_T(\sigma) = \frac{1}{Z(T)} e^{-E/T}$$

$$C = \mathrm{Var}_T(E/T)$$



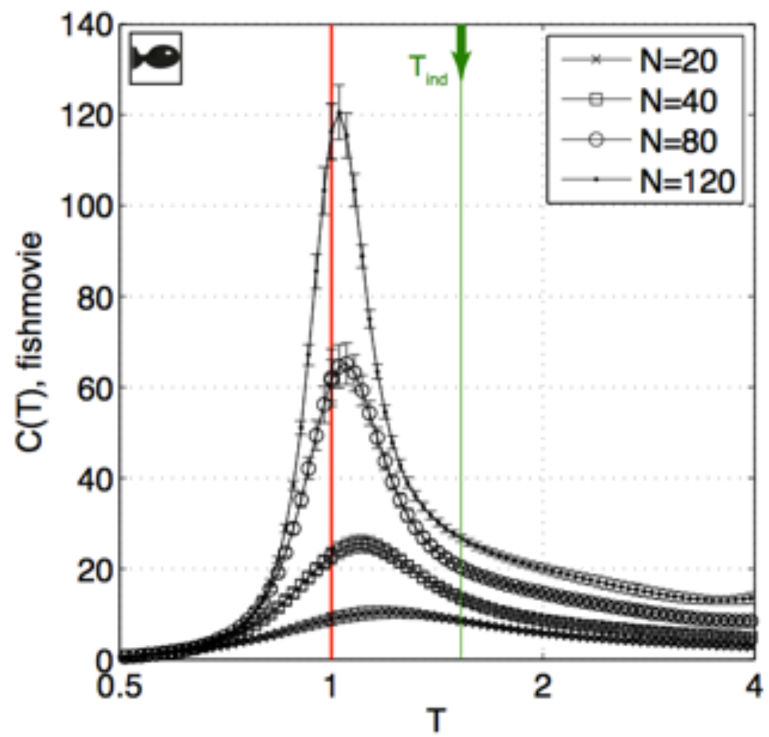


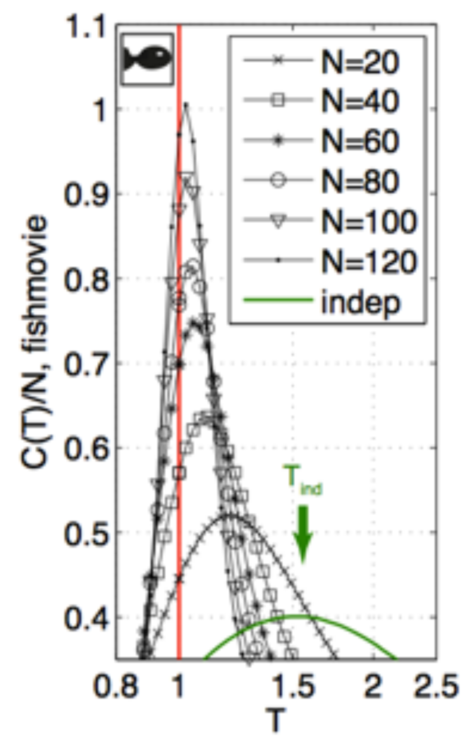
v = temporal range

Mora Deny Marre PRL 2015

thermodynamics of spike trains

static (salamander)



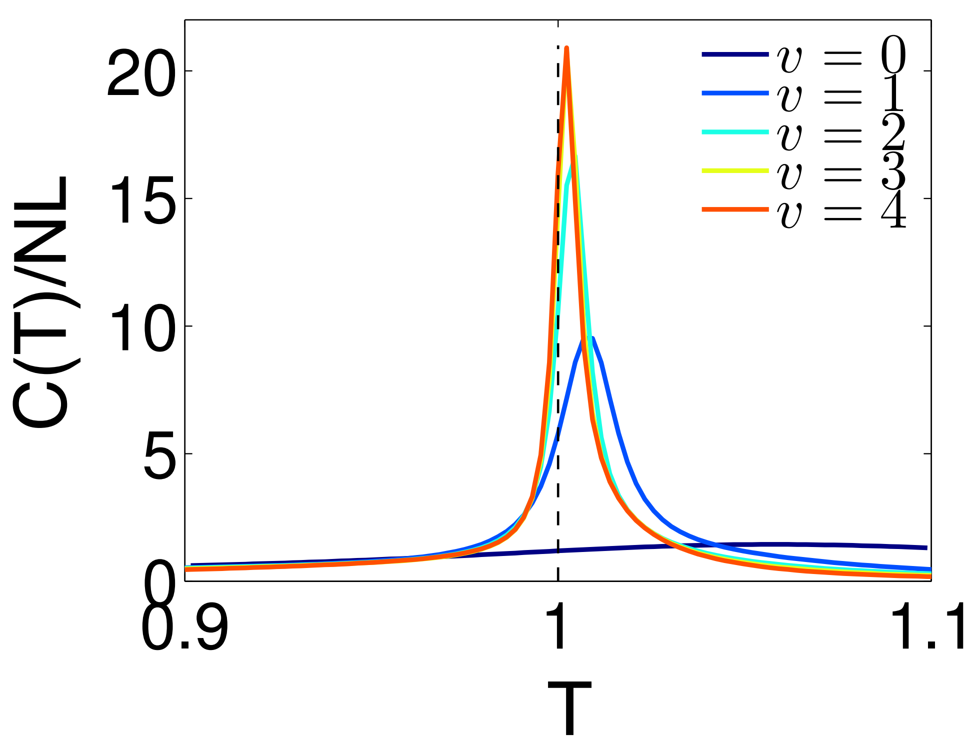


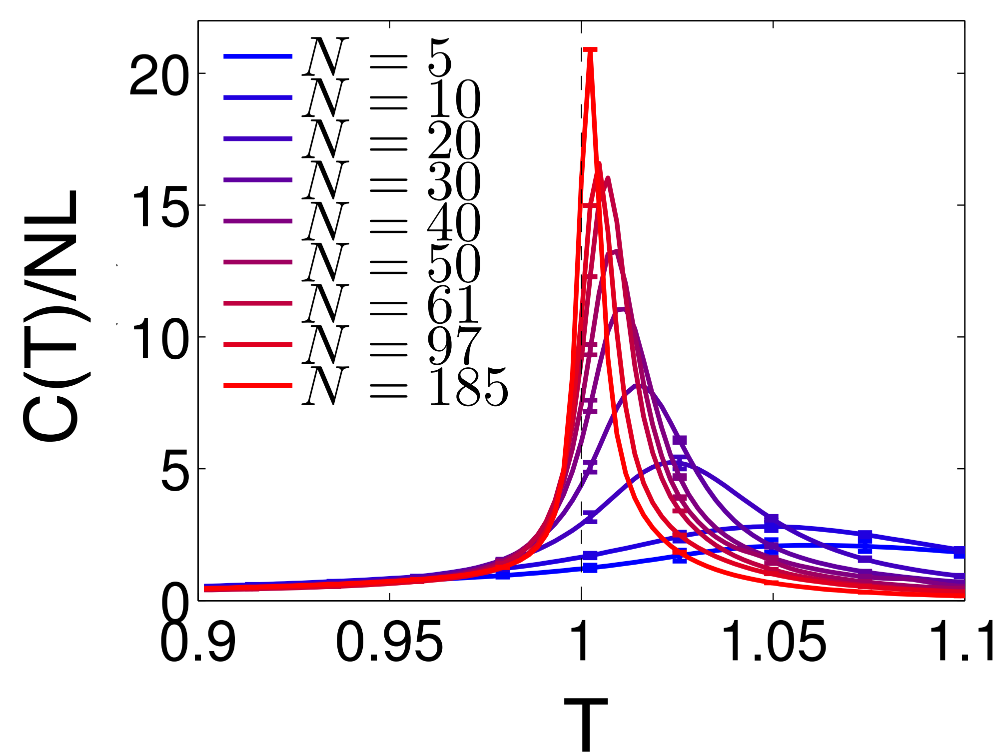
$$E = -\log P(\{\sigma_{i,t}\})$$

$$P_T(\sigma) = \frac{1}{Z(T)} e^{-E/T}$$

$$C = \mathrm{Var}_T(E/T)$$

dynamic (rat)

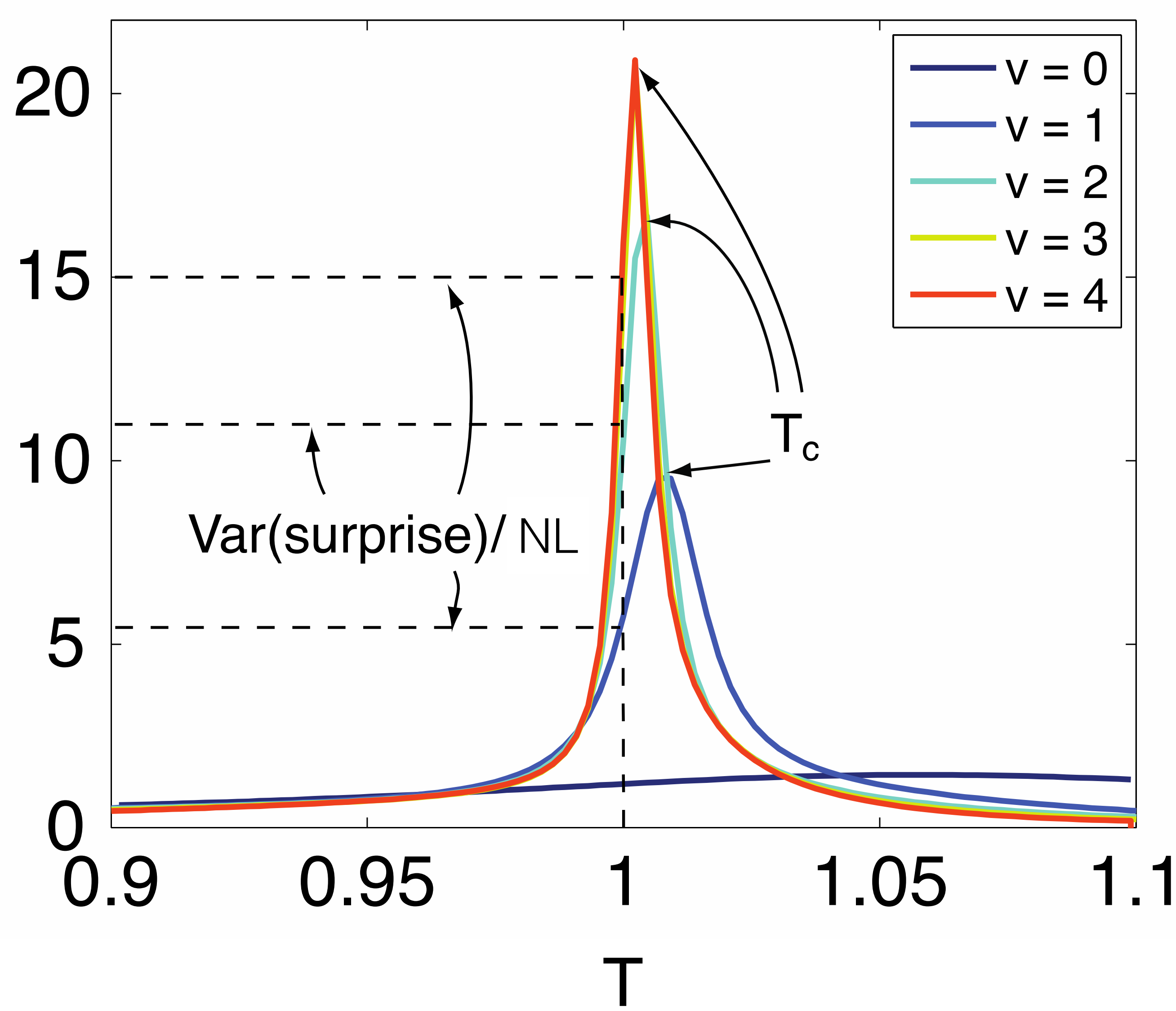


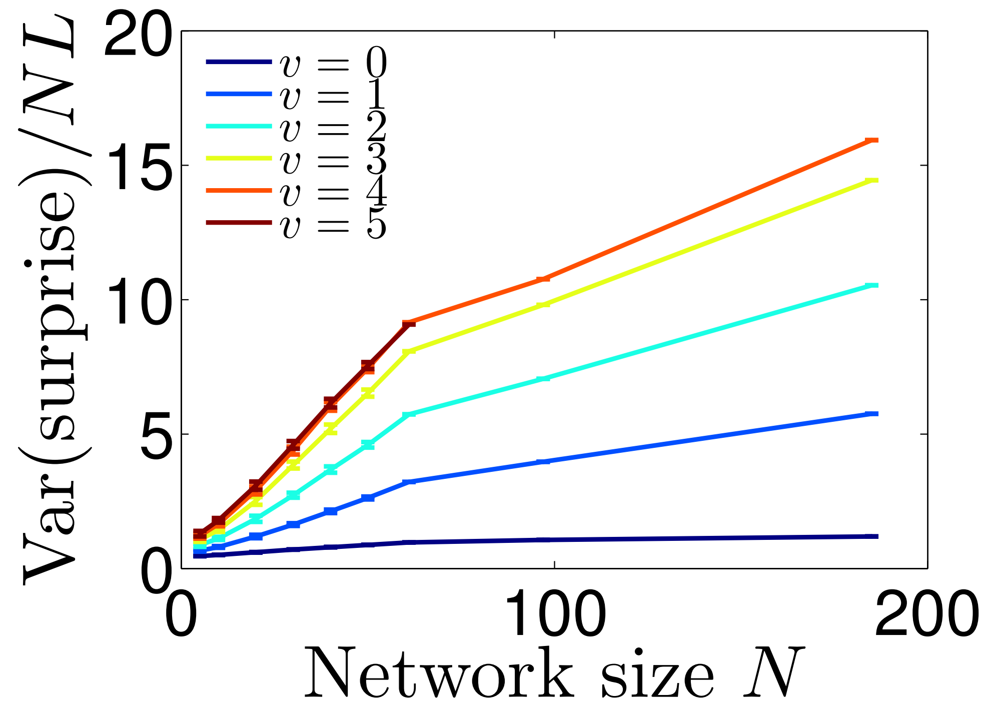


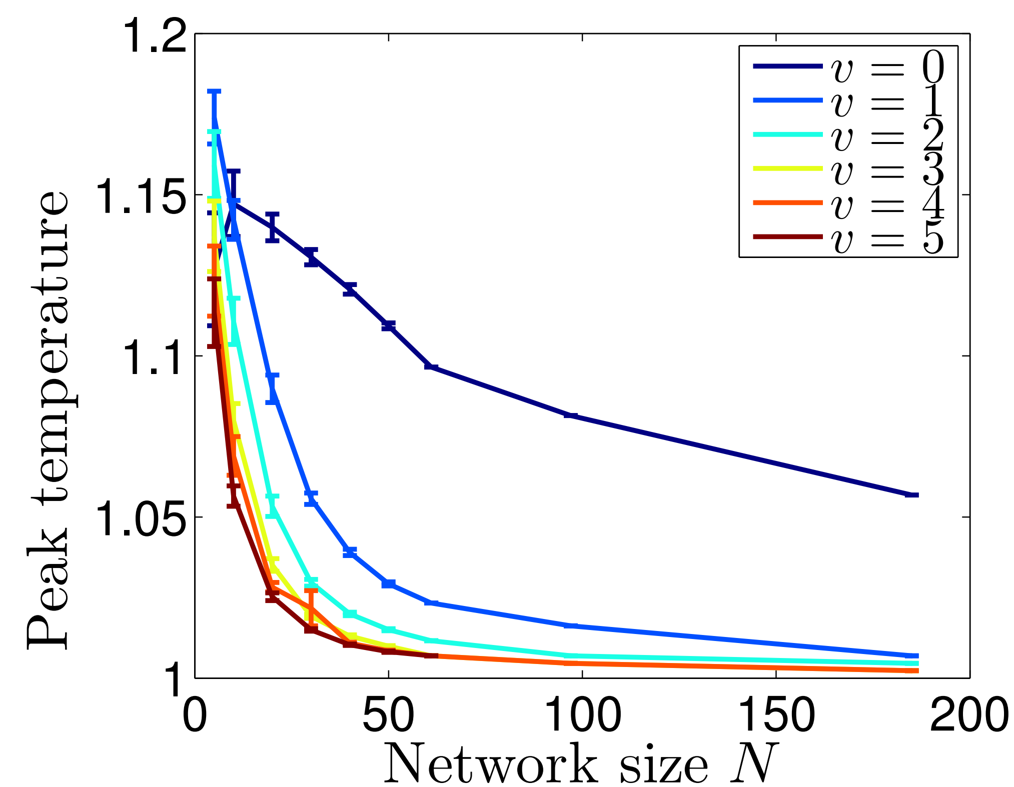
v = temporal range

Mora Deny Marre PRL 2015

scaling with network size



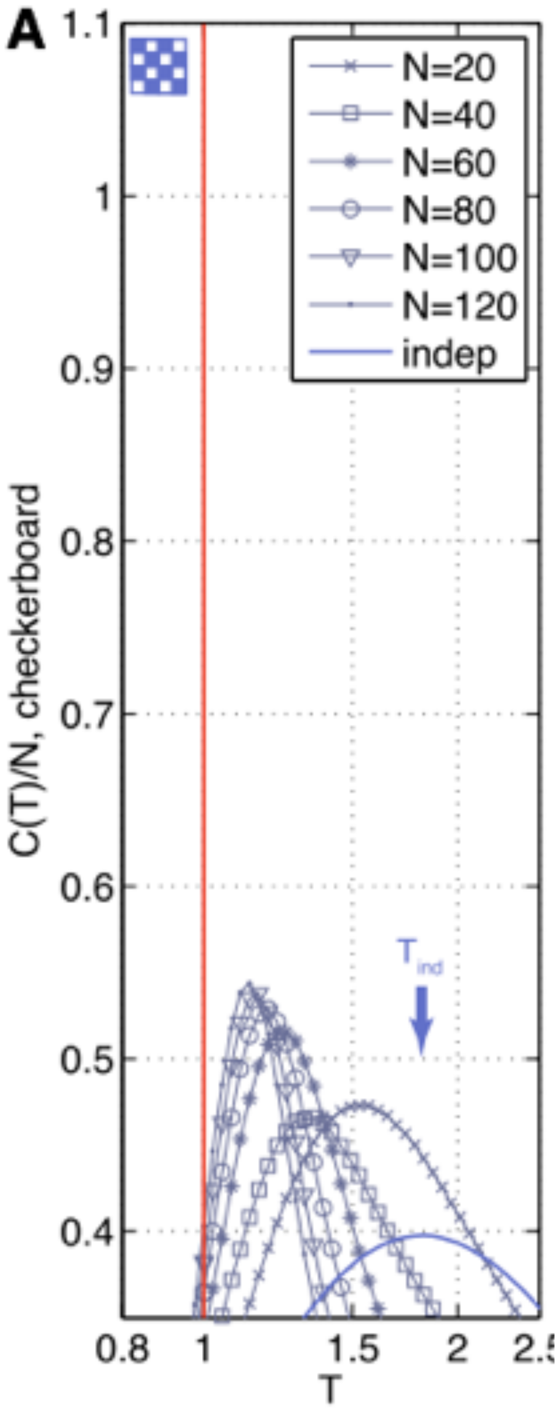


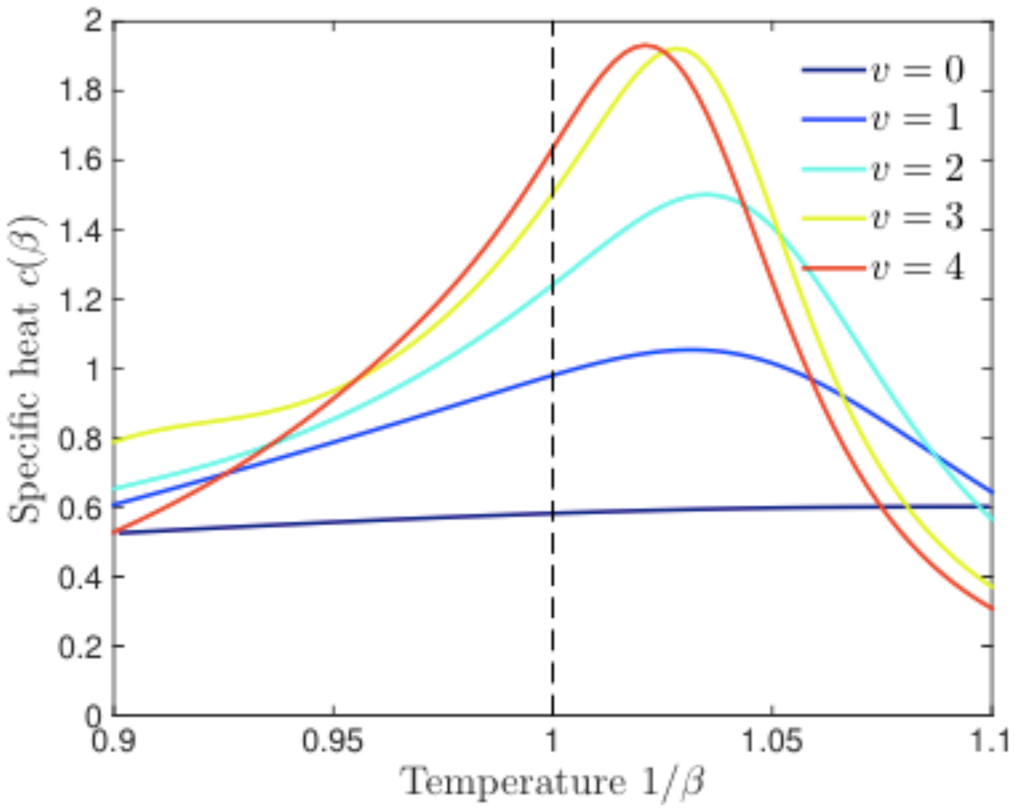


conclusions

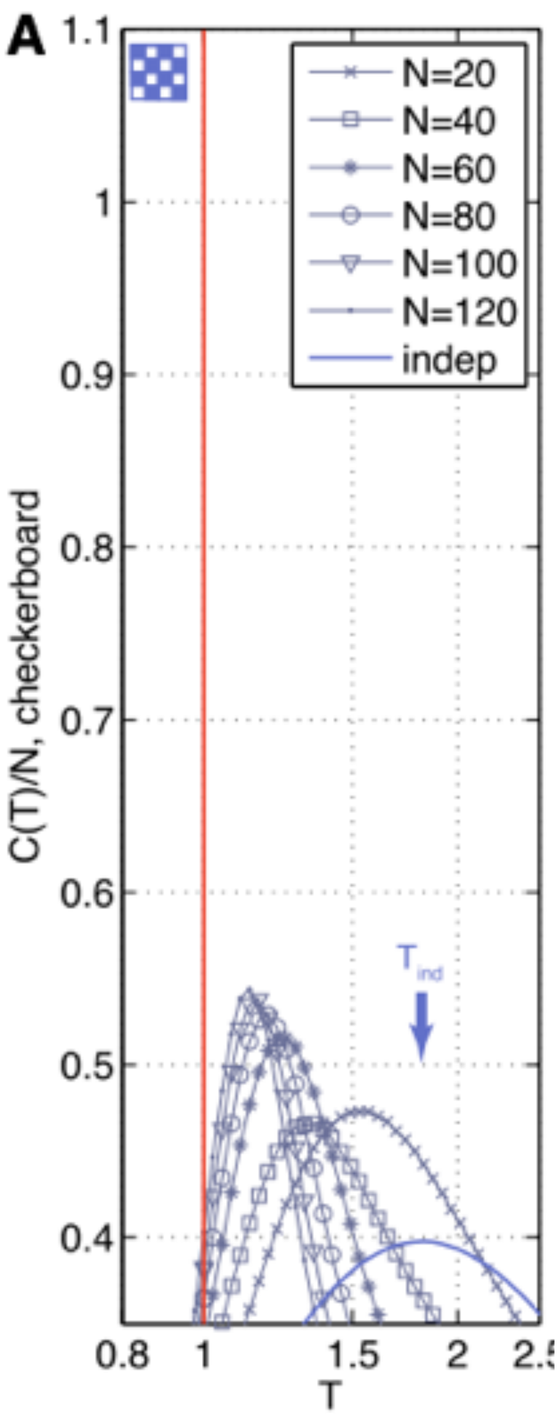
- stationary maximum entropy models may capture emergent behaviour in biological data
- but dynamic framework may be necessary to get parameters right
- in neural systems, heat capacity = useful indicator of critical properties
- critical signature enhanced by dynamical approach
- application to other biological contexts?

random flickering checkerboard

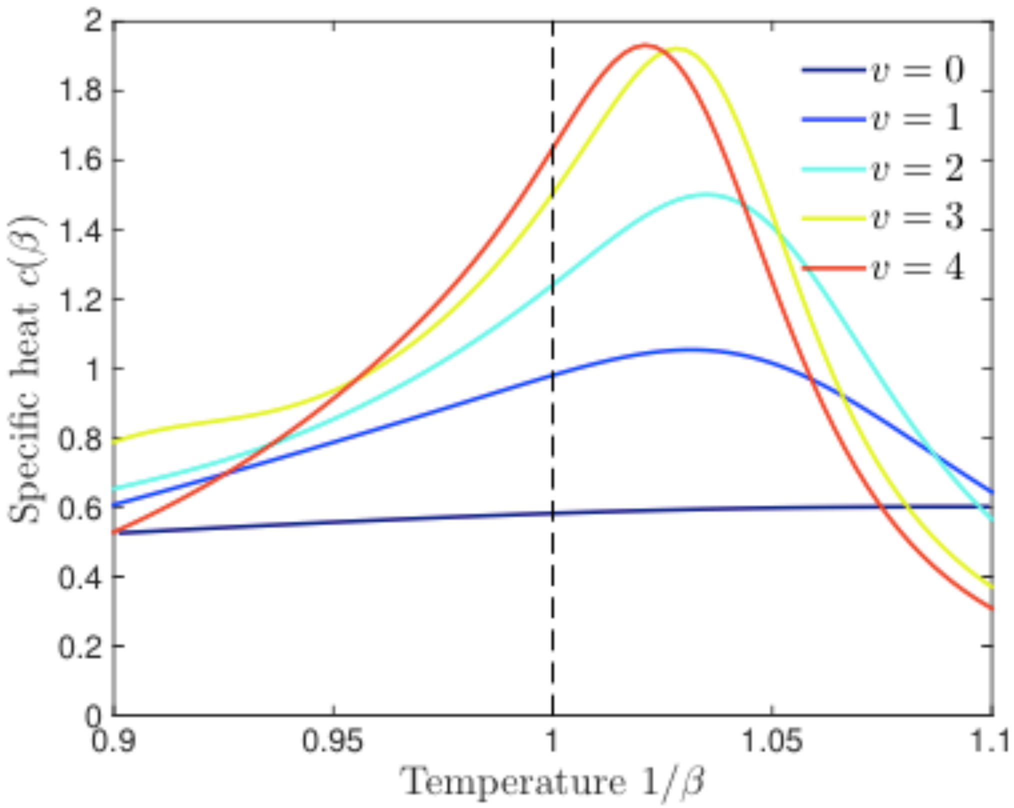




random flickering checkerboard



static
(salamander)



dynamic
(rat)