Spin glasses and proteins:

an old tool for an old problem ?

Olivier Rivoire

Laboratoire Interdisciplinaire de Physique
Grenoble, France

Cargèse - 2014

Two old problems

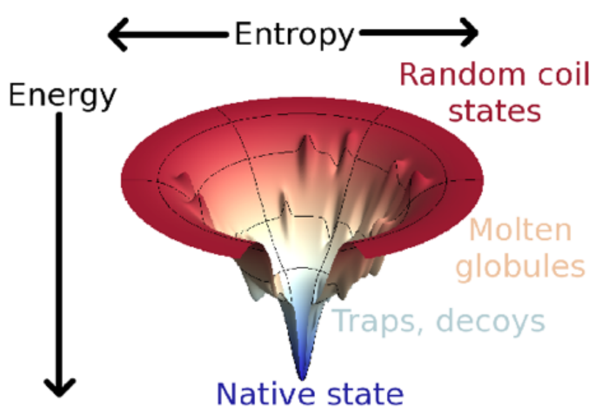
SEQUENCE $\xrightarrow{(1)}$ STRUCTURE $\xrightarrow{(2)}$ FUNCTION

(1) Folding

How do proteins find their 'native state' ?

Understood in principle (with ideas from spin glasses)

Funneled energy landscape / principle of minimum frustration



Bryngelson & Wolynes, *Spin glasses and the statistical mechanics of protein folding*, PNAS 1987

[Review] Dill et al., *The protein folding problem*, Annu. Rev. Biophysics 2008

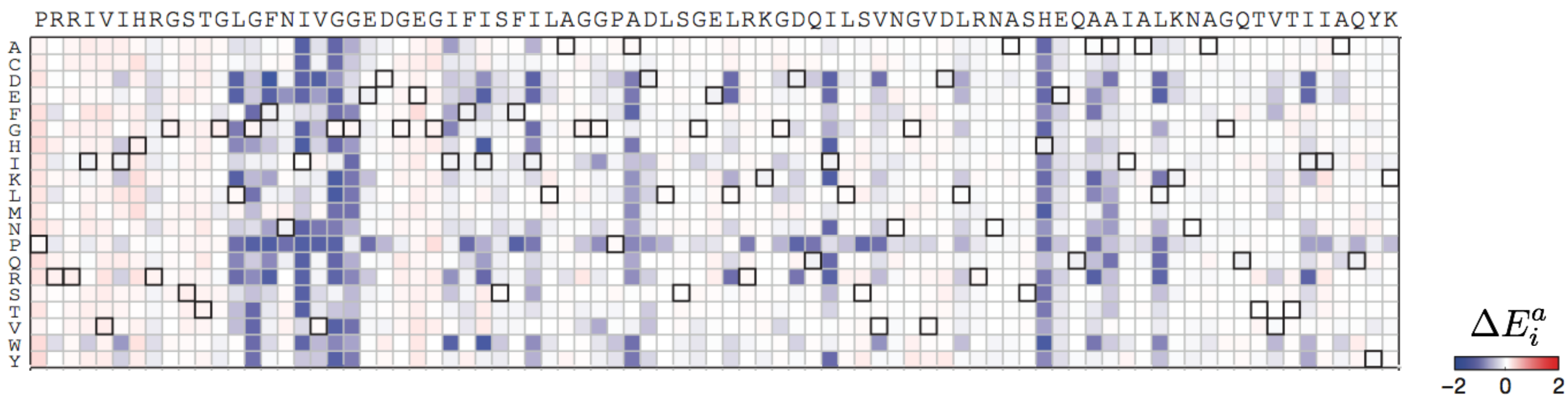
(2) Function

How do proteins work ?

New quantitative data

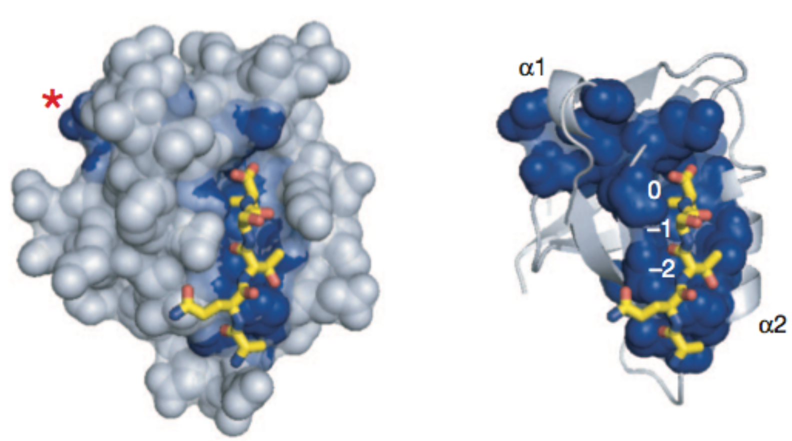
Saturated mutagenesis

Effects of point mutations on the binding affinity to a ligand (in PDZ):



McLaughlin et al., *The spatial architecture of protein function and adaptation*, Nature 2012

3d view:



How to understand this map ?

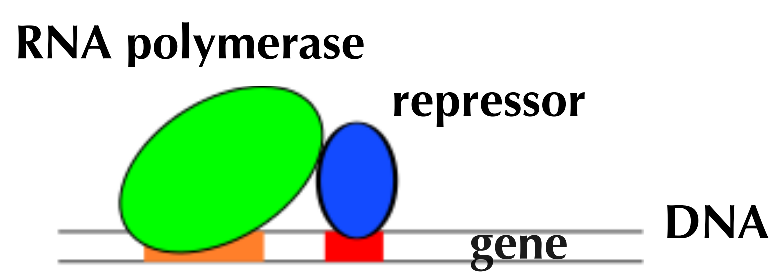
- Evolution: conservation and coevolution
- Experiments: allosteric effect

Allosteric regulation

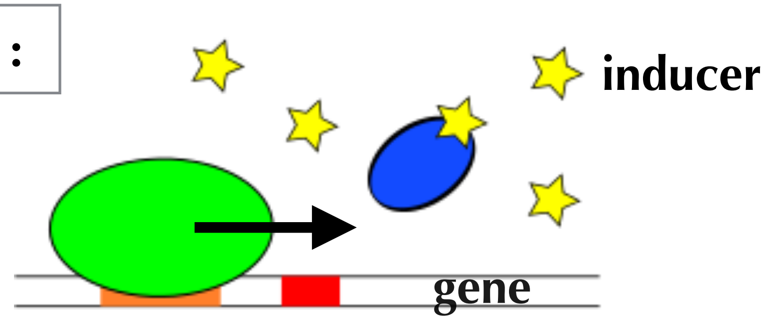
Definition: regulation of an 'active site' by a distant site on the protein

Repressor:

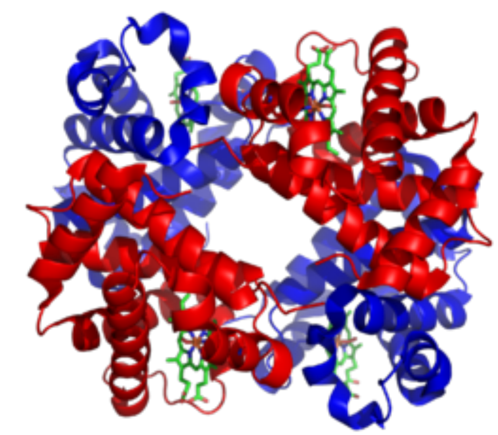
Repression of gene transcription :



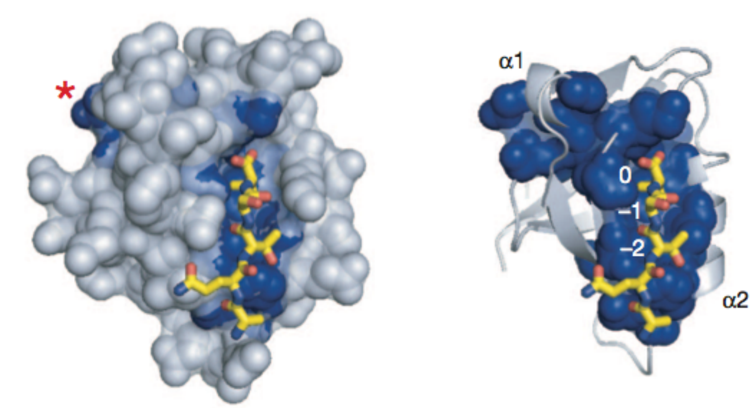
Activation :



Hemoglobin:



PDZ:

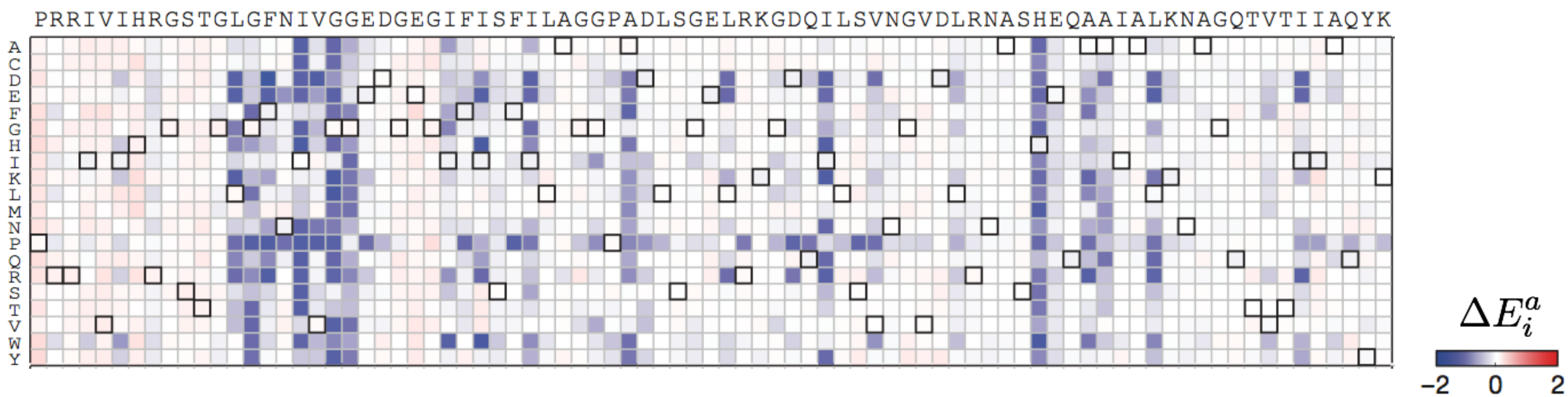


Monod, Changeux & Jacob, *Allosteric proteins and cellular control systems*, J. Mol. Biol. 1967

[Review] Smock & Gierasch, *Sending messages dynamically*, Science 2009

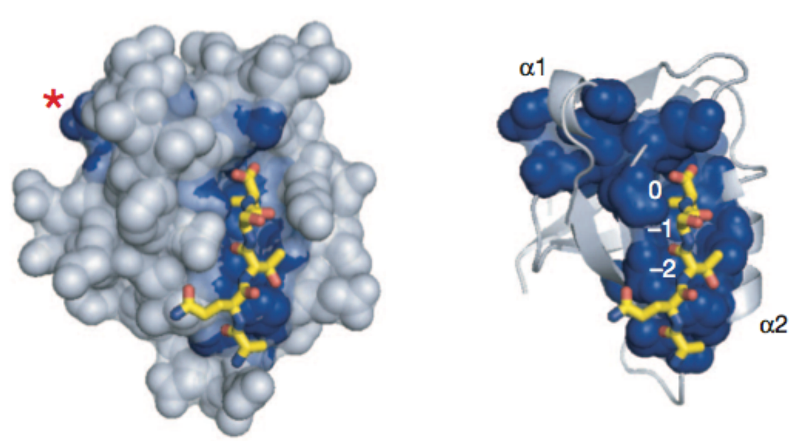
How to understand this “structure” ?

Effects of point mutations on the binding affinity to a ligand (in PDZ):



McLaughlin et al., *The spatial architecture of protein function and adaptation*, Nature 2012

3d view:



- Evolution: conservation and coevolution
- Experiments: allosteric effect

It need NOT be PHYSICS, it may be EVOLUTION

A spin glass model for the evolution of allostery

A spin glass model (1): binding

$$E[\sigma|a,\varepsilon] = -\sum_{\langle i,j\rangle} J(a_i,a_j)\sigma_i\sigma_j - \sum_i h(\varepsilon_i)\sigma_i$$

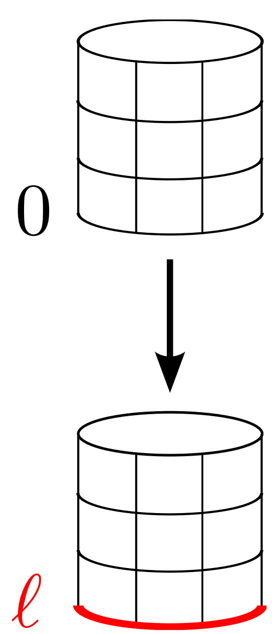
fixed structure (lattice)

amino acid

physical state (ex: orientation)

environment (ex: ligand)

3 types of variables / 3 time scales



Solvent: $h_i = 0$

Ligand: $h_i = \ell_i \quad (i \in \text{ bottom})$

Binding free energy:

$$\Delta F(a,\ell) = F(a,h=\ell) - F(a,h=0)$$

A spin glass model (2): allostery

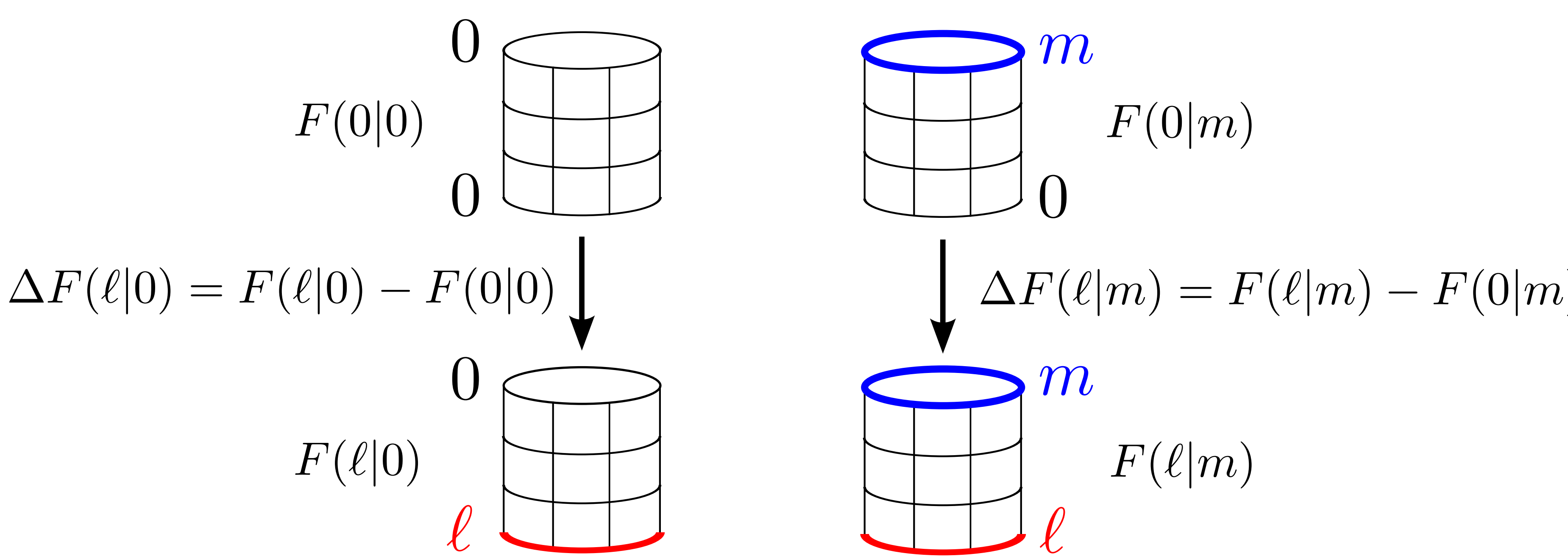
$$E[\sigma|a,\varepsilon] = -\sum_{\langle i,j\rangle} J(a_i,a_j)\sigma_i\sigma_j - \sum_i h(\varepsilon_i)\sigma_i$$

fixed structure (lattice)

amino acid

physical state (ex: orientation)

environment (ex: ligand)



Allosteric efficiency: $\phi(m,\ell) = \Delta F(\ell|0) - \Delta F(\ell|m)$

= how much better ℓ binds in the presence of m

Evolving spin glasses (1): principles

Population dynamics:

selection / reproduction

$$\{J_{ij}^{(1)}\}_{ij} \quad \vdots \quad \{J_{ij}^{(k)}\}_{ij} \quad \vdots \quad \{J_{ij}^{(P)}\}_{ij}$$

$$\left.\begin{array}{c}\{J_{ij}^{(k)}\}_{ij}\\ \{J_{ij}^{(k)}\}_{ij}\\ \{J_{ij}^{(k)}\}_{ij}\end{array}\right\} n^{(k)} = f\left(\phi(J^{(k)})\right)$$

allosteric efficiency

$$\{J_{ij}^{(1)}\}_{ij} \quad \vdots \quad \{J_{ij}^{(k)}\}_{ij} \quad \vdots \quad \{J_{ij}^{(P)}\}_{ij}$$

random mutations

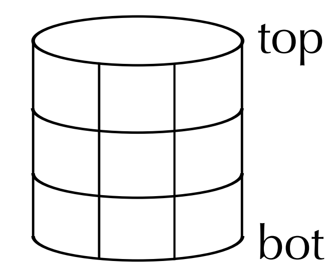
$$J_{ij}^{(k)} \leftarrow \text{r.v.} \qquad \text{proba } \mu$$

$$J_{ij}^{(k)} \leftarrow J_{ij}^{(k)} \qquad \text{proba } 1-\mu$$

[Book] Mitchell, *An introduction to genetic algorithms*, 1998

Evolving spin glasses (2): technicalities

- **Gaussian spin glass:** $Z\,(h|J) = \int \prod_i \frac{d\sigma_i \; e^{-\sigma_i^2/2}}{\sqrt{2\pi}} e^{\beta\left(\frac{1}{2}\sigma^\top J\sigma - h^\top \sigma\right)}$



- **Allosteric efficiency:** $\phi(J|m,\ell) = \beta\; m^\top[(I-\beta J)^{-1}]_{\mathrm{top,bot}}\;\ell$

- **Mutations:** $J^*_{ij} \sim U([-1,+1])$ (uniform distribution)

  $J^{(k)}_{ij} \leftarrow \text{r.v.}\quad \text{proba } \mu$

  $J^{(k)}_{ij} \leftarrow J^{(k)}_{ij}\quad \text{proba } 1-\mu$

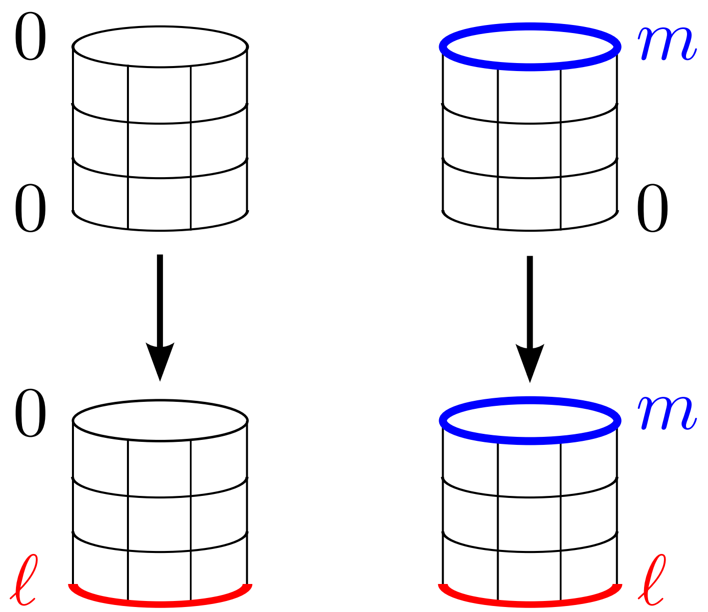
- **Parameters:** $N = 10 \times 10$ (system size) $\quad \beta = 0.1$ (temperature)

  $P = 500$ (population size) $\quad \mu = 10^{-5}$ (mutation rate)

- **Sigma-scaling rule:** $n_k = f(\phi_k) = 1 + \frac{\phi_k - \overline{\phi}}{2\sigma_\phi^2}$

$$\{J^{(k)}_{ij}\}_{ij} \rightarrow \left.\begin{array}{l}\{J^{(k)}_{ij}\}_{ij}\\ \{J^{(k)}_{ij}\}_{ij}\\ \{J^{(k)}_{ij}\}_{ij}\end{array}\right\}\; n^{(k)} = f\left(\phi(J^{(k)})\right)$$

Evolving spin glasses (3): results



$\ell = (+1, \ldots, +1)$

$m = (+1, \ldots, +1)$

Energy:

$$E[\sigma | J, m, \ell] = -\sum_{\langle i,j \rangle} J_{ij}\sigma_i\sigma_j - \sum_{i \in \mathrm{top}} m_i\sigma_i - \sum_{i \in \mathrm{bot}} \ell_i\sigma_i$$

$$J_{ij} \in [-1, 1]$$

Fitness:

$$\phi(m, \ell) = F(\ell, 0) - F(0, 0) - F(\ell, m) + F(0, m)$$

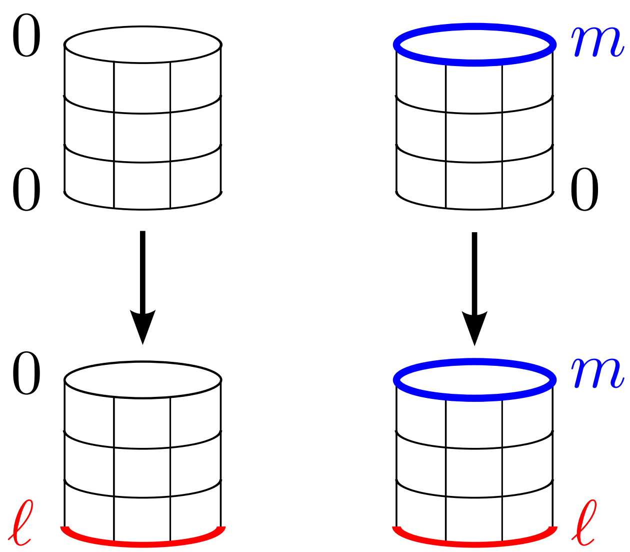
Outcome of evolution:

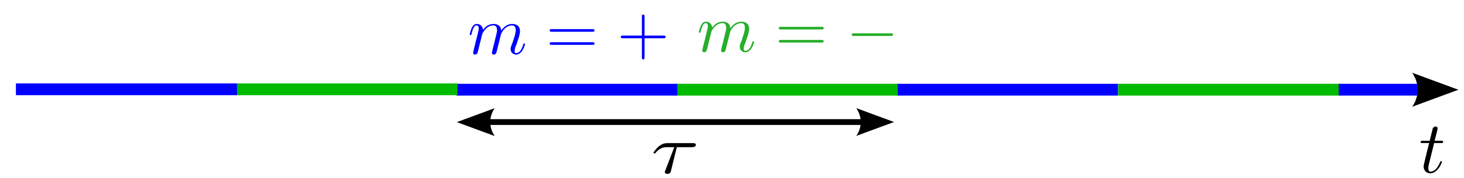
$$\boxed{|J_{ij}| \simeq 1 \quad \forall\ i, j}$$

Interpretation: ferromagnet with maximal couplings

no sparse architecture !

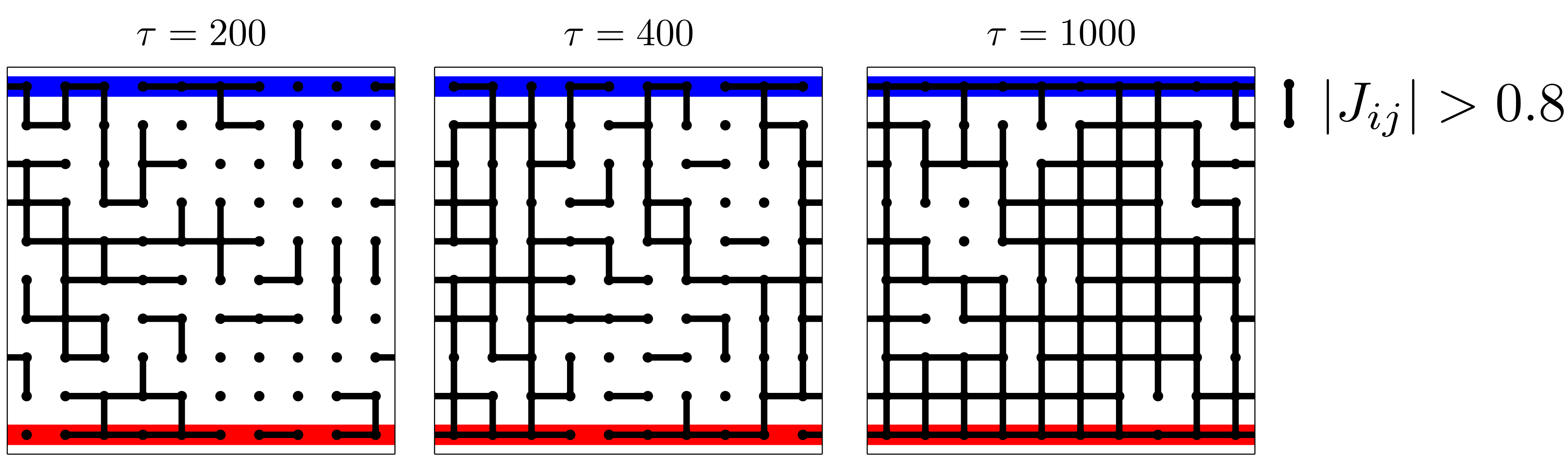
Sparsity from fluctuations (1)





$\ell = (+1, \ldots, +1)$

$$\left.\begin{aligned} m &= (+1, \ldots, +1) \\ m &= (-1, \ldots, -1) \end{aligned}\right\}$$ change every $\tau/2$ generations



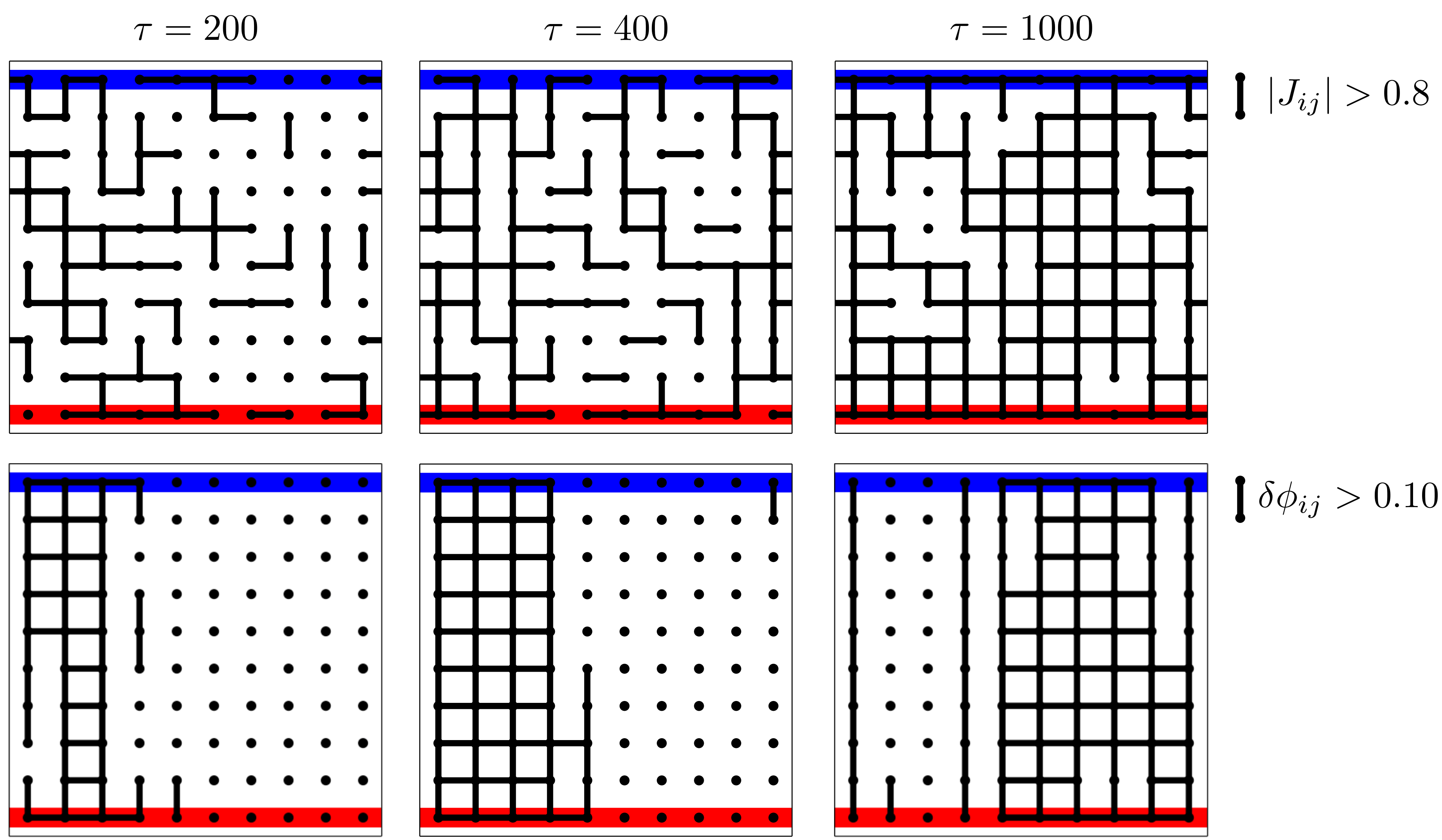
More fluctuations → more sparsity

Sparsity from fluctuations (2)

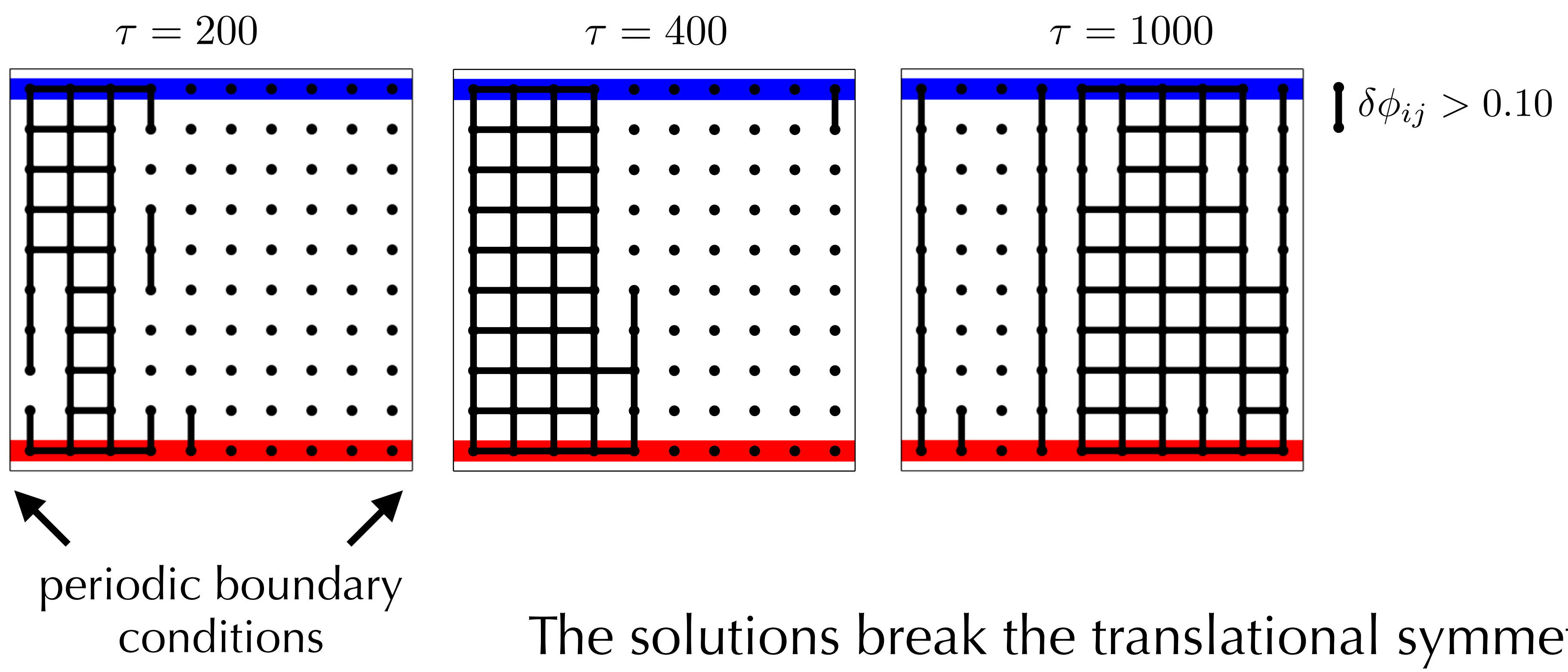
Functional value : $\delta\phi_{ij} =$ fitness cost for mutating J_{ij}

large couplings

large fitness costs

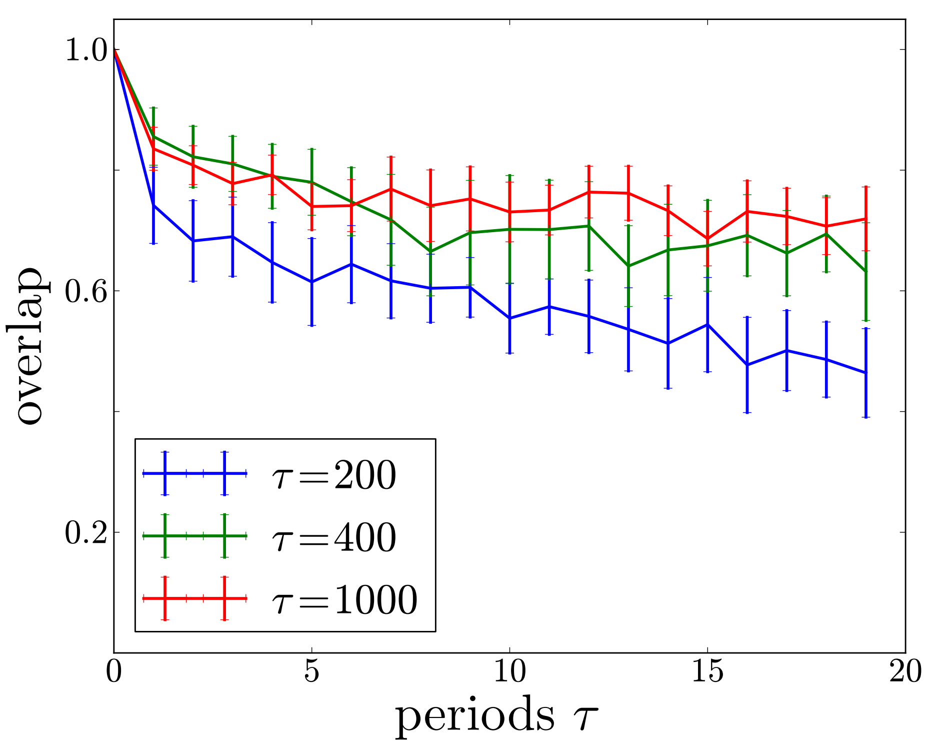


Evolutionary conservation



The solutions break the translational symmetry

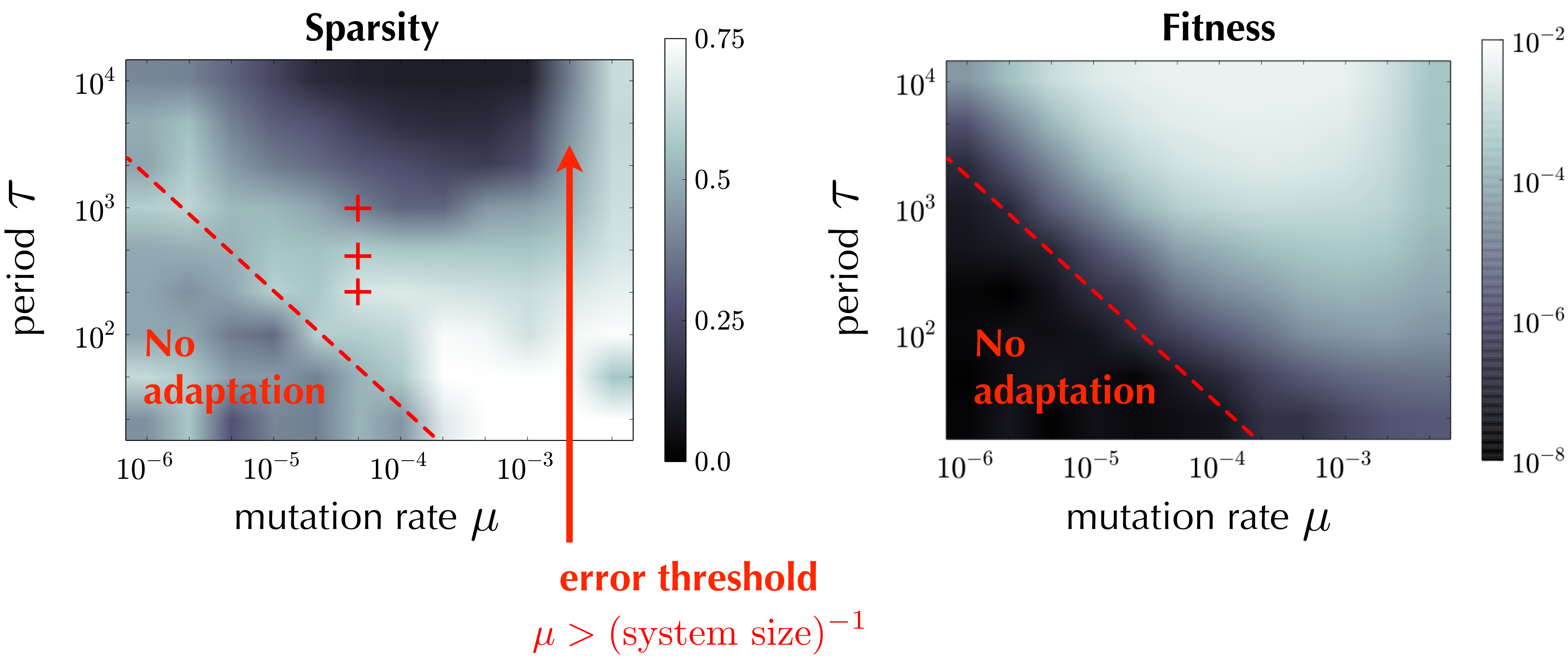
The location of the channel is stable over multiple periods :



Phase diagram

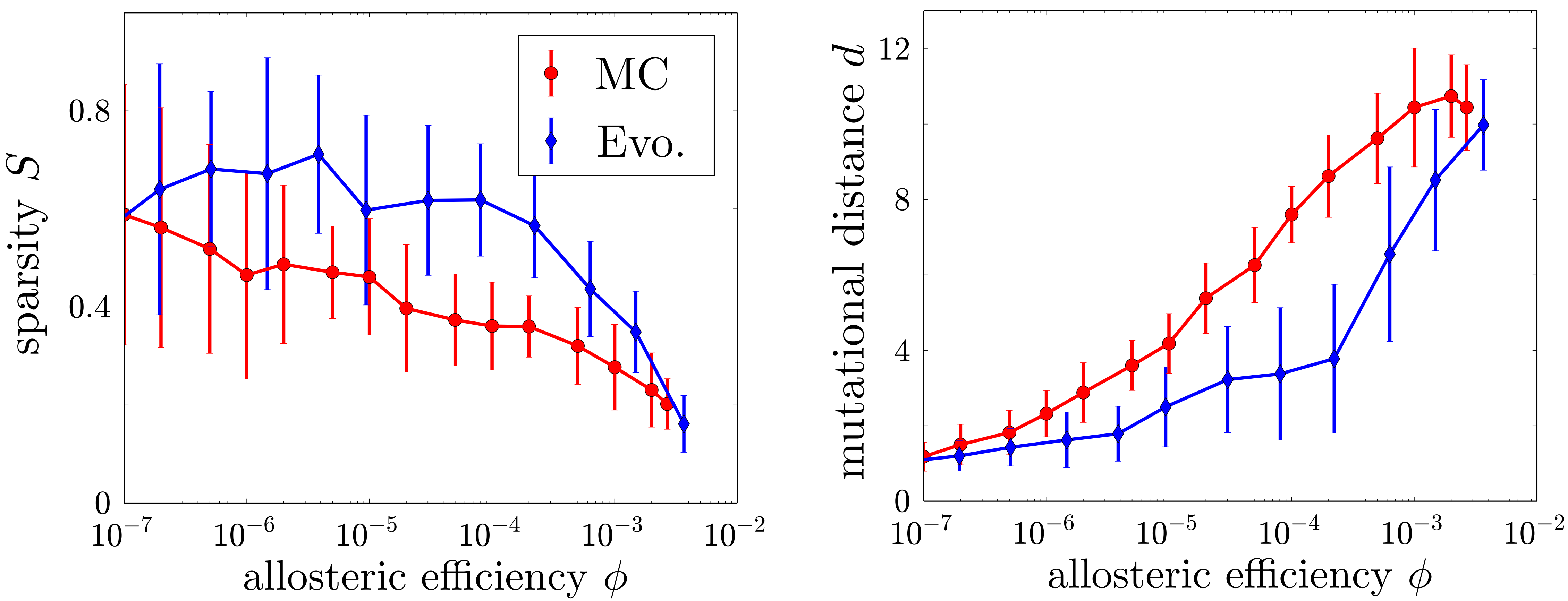
Sparsity: fraction of couplings J_{ij} with small fitness cost upon mutation $(\delta\phi_{ij} < 0.1)$

Fitness: allosteric efficiency ϕ



Two routes to sparsity:
- fluctuating selective pressures (intermediate $\mu\tau$)
- mutational load (large μ)

Sparsity beyond low fitness



Systems evolved with fluctuating selection are:

- more sparse
- more 'evolvable'

than systems with equivalent fitness ϕ

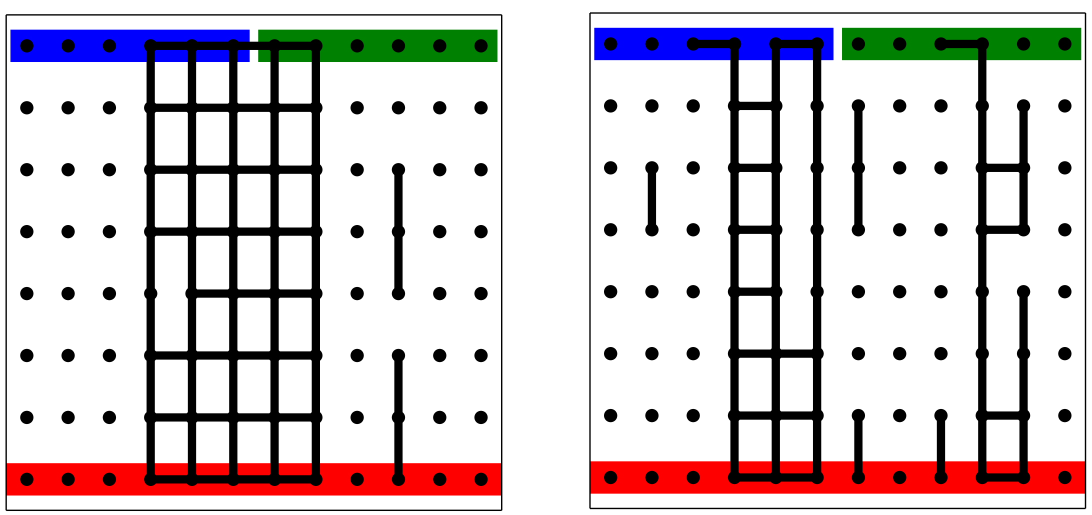
Geometry beyond sparsity

temporal structure
of selective pressures
in past history

→

spatial structure
of couplings
in present systems

Fitness: binding of ℓ
if m_1 or/and m_2 are present



Probability of modularity:

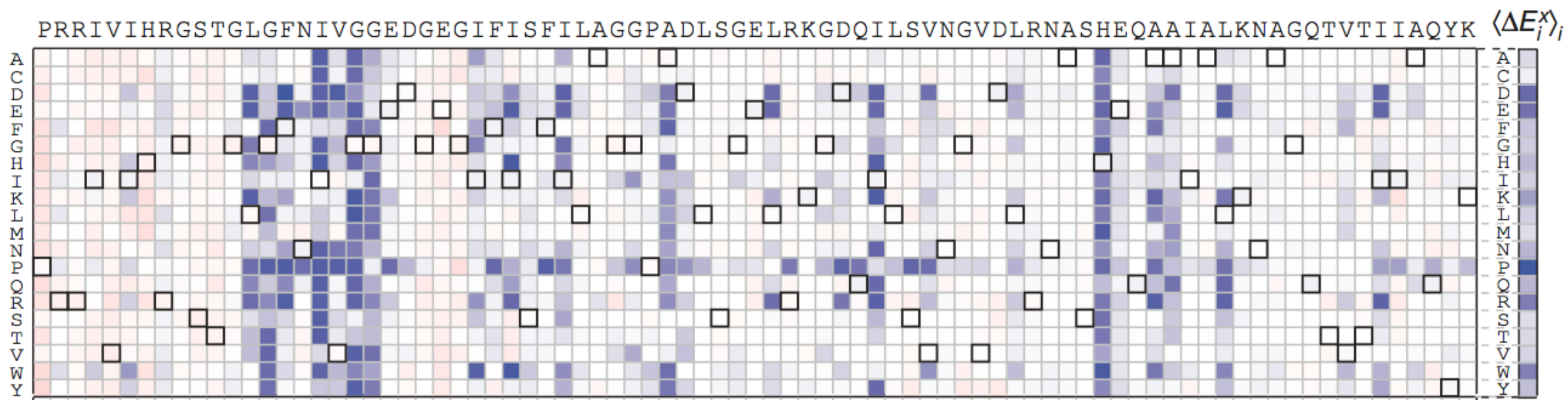
	$\tau = 100$	$\tau = 200$
m_1, m_2 vary together	0.03	0.10
m_1, m_2 vary one at a time	0.82	0.51

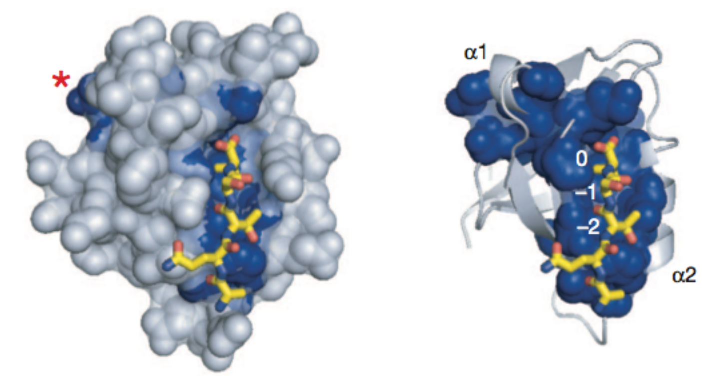
more likely to be modular when
m_1 and m_2 vary independently

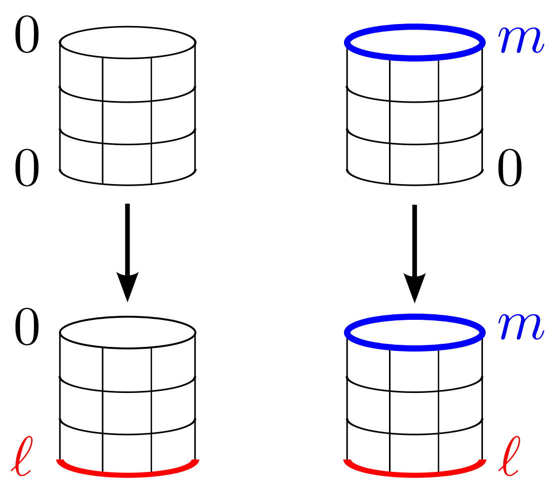
modularly varying environment → **modular geometry**

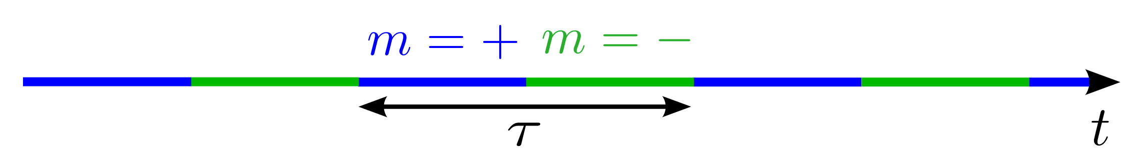
[similar results] Kashtan & Alon, *Spontaneous evolution of modularity and network motifs*, PNAS 2005

How a “functional structure” may be shaped by temporal fluctuations

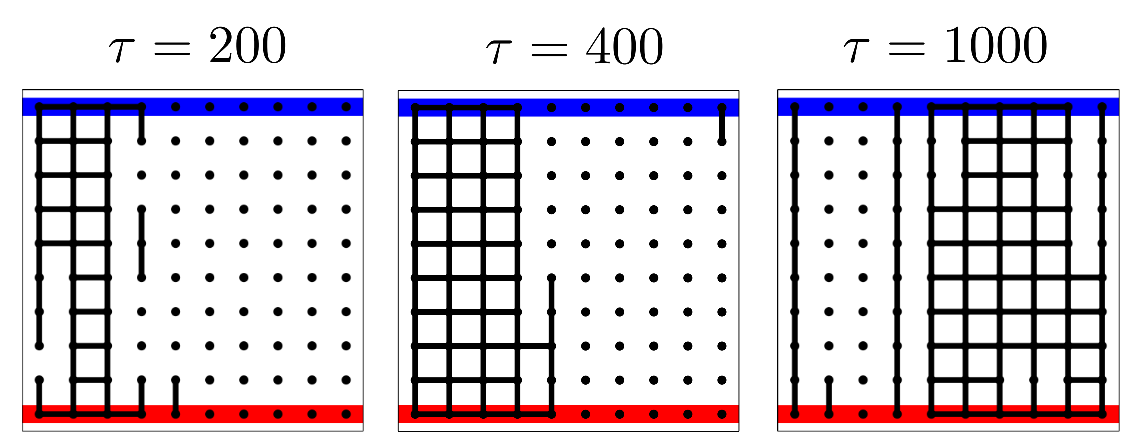








$$E[\sigma|J,m,\ell] = -\sum_{\langle i,j \rangle} J_{ij}\sigma_i\sigma_j - \sum_{i \in \mathrm{top}} m_i\sigma_i - \sum_{i \in \mathrm{bot}} \ell_i\sigma_i$$



M. Hemery & O. Rivoire, *Evolution of sparsity and modularity in a model of protein allostery,* arXiv:1408.3280