# Evolution and Replicas

Yerevan Workshop
Evolution of Complexity from the Statistical Physics Perspective

ACC Coolen

Radboud University, Nijmegen
June 2022

1. Introduction
   - Biological evolution as a modelling challenge
   - Heterogeneity in physics versus biology
2. Inspiration from statistical physics
   - Ising spin systems with Hebbian bond evolution
   - XY spins and added layers of complexity
   - Link with the Parisi scheme
3. Application to biology
   - Self-programming in neural systems
   - Genesis of protein structure
   - Finite-*n* replica approach to biological evolution
4. Open questions
5. Summary

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-$n$ replica approach to biological evolution

4 Open questions

5 Summary

# Biological evolution as a modelling challenge

many-particle physics
versus many-particle biology ...

*in common*

- many stochastically evolving variables
- many interactions, via nonlinear equations

*key differences*

- physics: usually detailed balance processes,
  evolution towards equilibrium state, $p_\infty(\boldsymbol{\sigma}) = Z^{-1} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})}$

  biology: almost never detailed balance,
  we don't know the stationary state, $p_\infty(\boldsymbol{\sigma}) = ?$

- the nature of heterogeneity ...

# Biological evolution as a modelling challenge

many-particle physics
versus many-particle biology ...

*in common*

- many stochastically evolving variables
- many interactions, via nonlinear equations

*key differences*

- physics: usually detailed balance processes,
  evolution towards equilibrium state, $p_\infty(\boldsymbol{\sigma}) = Z^{-1} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})}$

  biology: almost never detailed balance,
  we don't know the stationary state, $p_\infty(\boldsymbol{\sigma}) = ?$

- the nature of heterogeneity ...

# Biological evolution as a modelling challenge

many-particle physics
versus many-particle biology ...

*in common*

- many stochastically evolving variables
- many interactions, via nonlinear equations

*key differences*

- physics: usually detailed balance processes,
  evolution towards equilibrium state, $p_\infty(\boldsymbol{\sigma}) = Z^{-1}\mathrm{e}^{-\beta H(\boldsymbol{\sigma})}$

  biology: almost never detailed balance,
  we don't know the stationary state, $p_\infty(\boldsymbol{\sigma}) = ?$
- the nature of heterogeneity ...

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

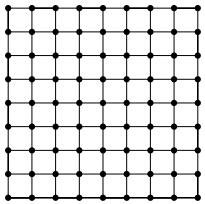
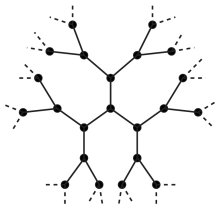
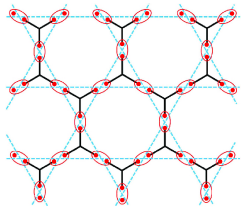
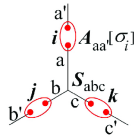
3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-*n* replica approach to biological evolution

4 Open questions

5 Summary

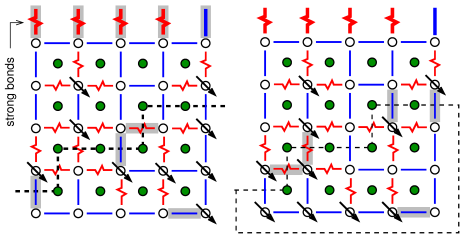
# The nature of heterogeneity

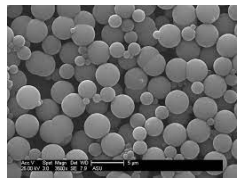
- most stat phys methods are designed for *homogeneous* systems and *regular* topologies


  – particles interchangeable
  – either uniform ‘all-to-all’ interactions (mean-field), or
  – use symmetries (transfer matrices, renormalization, ...)

- statistical physics of *heterogeneous* systems





  – heterogeneity is *random* (random forces, particle sizes, topologies, ...)
  – use disorder-averaged generating functions (replica analysis, generating functional analysis, ...)
  – microscopic realization of disorder irrelevant

# Analysis of heterogeneous systems in physics

map to a (more tricky) *homogeneous* system

- **Replica method**

  compute disorder-averaged free energy density

$$\begin{aligned} \overline{f} &= -\lim_{N\to\infty}\frac{1}{\beta N}\overline{\log Z} = -\lim_{N\to\infty}\frac{1}{\beta N}\lim_{n\to 0}\frac{1}{n}\log\overline{\left[\sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})}\right]^n} \\ &= -\lim_{n\to 0}\frac{1}{\beta n}\lim_{N\to\infty}\frac{1}{N}\log\left[\sum_{\boldsymbol{\sigma}^1\ldots\boldsymbol{\sigma}^n}\overline{\mathrm{e}^{-\beta\sum_{\alpha=1}^n H(\boldsymbol{\sigma}^\alpha)}}\right] = \ldots \end{aligned}$$

  final result:
  single particle eqn in equilibrium, $n$-replicated with $n\to 0$ ...

- **Generating functional analysis**

  compute disorder-average generating functional

$$\overline{Z[\psi]} = \overline{\mathrm{e}^{\mathrm{i}\sum_{ti}\psi_i(t)\sigma_i(t)}} = \sum_{\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t_{\max})} \mathrm{e}^{\mathrm{i}\sum_{ti}\psi_i(t)\sigma_i(t)}\,\overline{P(\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t_{\max}))} = \ldots$$

  final result:
  single particle eqn, retarded self-interaction + non-white noise ...

# Analysis of heterogeneous systems in physics

map to a (more tricky) *homogeneous* system

- **Replica method**

  compute disorder-averaged free energy density

$$\begin{aligned} \overline{f} &= -\lim_{N\to\infty}\frac{1}{\beta N}\overline{\log Z} = -\lim_{N\to\infty}\frac{1}{\beta N}\lim_{n\to 0}\frac{1}{n}\log\overline{\left[\sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})}\right]^n} \\ &= -\lim_{n\to 0}\frac{1}{\beta n}\lim_{N\to\infty}\frac{1}{N}\log\left[\sum_{\boldsymbol{\sigma}^1\ldots\boldsymbol{\sigma}^n}\overline{\mathrm{e}^{-\beta\sum_{\alpha=1}^n H(\boldsymbol{\sigma}^\alpha)}}\right] = \ldots \end{aligned}$$

  final result:
  single particle eqn in equilibrium, $n$-replicated with $n\to 0$ ...

- **Generating functional analysis**

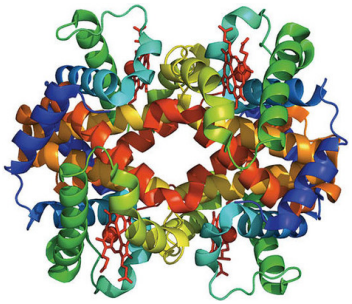
  compute disorder-average generating functional

$$\overline{Z[\psi]} = \overline{\mathrm{e}^{\mathrm{i}\sum_{ti}\psi_i(t)\sigma_i(t)}} = \sum_{\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t_{\max})} \mathrm{e}^{\mathrm{i}\sum_{ti}\psi_i(t)\sigma_i(t)}\,\overline{P(\boldsymbol{\sigma}(0),\ldots,\boldsymbol{\sigma}(t_{\max}))} = \ldots$$

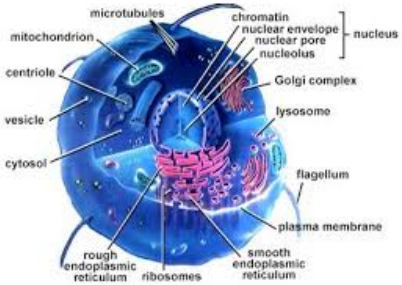
  final result:
  single particle eqn, retarded self-interaction + non-white noise ...

# Heterogeneity in biology: is *never* random ...

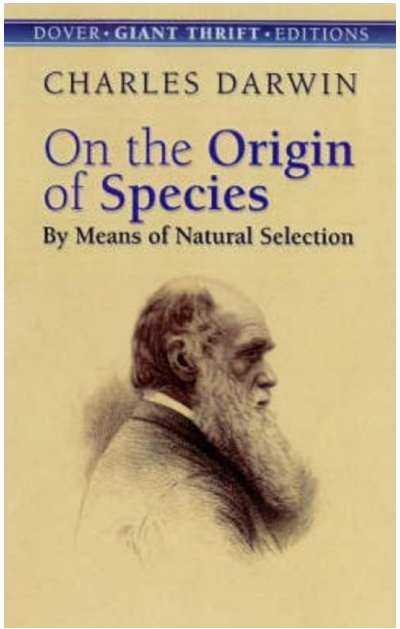
*molecular level*



*cellular level*




*organism level*

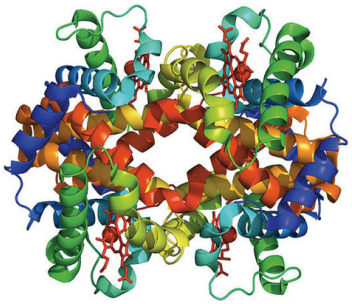



- heterogeneous parameters are *selected*, based on complex criteria ...
- microscopic realization of heterogeneity can be highly relevant
- we cannot average generating functions over the heterogeneity, because we cannot capture it in a probability distribution ...

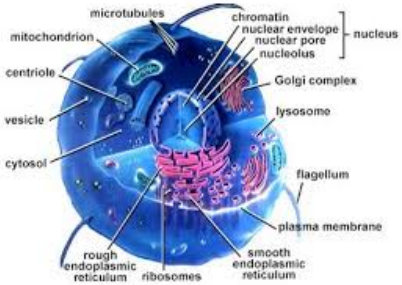
should we forget about disordered systems tools?
(replicas, generating functionals, cavity method, ...)

# Heterogeneity in biology: is *never* random ...

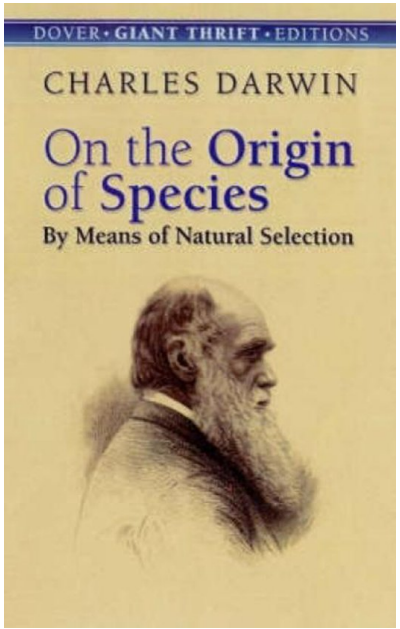
*molecular level*



*cellular level*




*organism level*

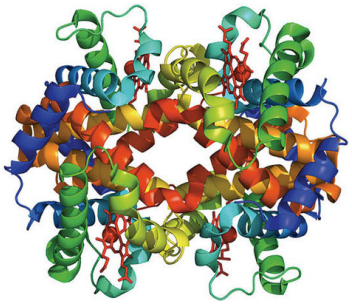



- heterogeneous parameters are *selected*, based on complex criteria ...
- microscopic realization of heterogeneity can be highly relevant
- we cannot average generating functions over the heterogeneity, because we cannot capture it in a probability distribution ...

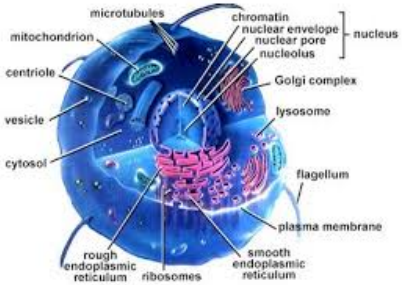
should we forget about disordered systems tools?
(replicas, generating functionals, cavity method, ...)

# Heterogeneity in biology: is *never* random ...

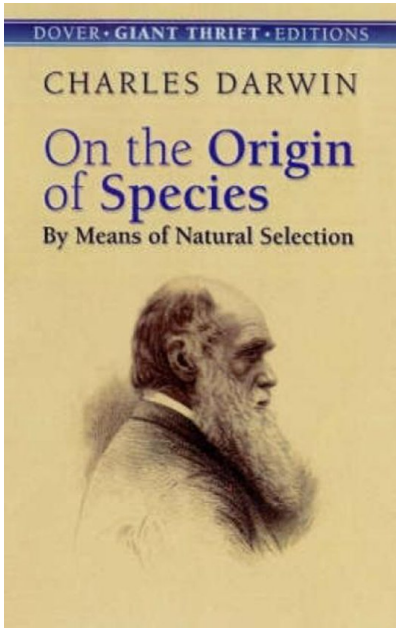
*molecular level*



*cellular level*




*organism level*




- heterogeneous parameters are *selected*, based on complex criteria ...
- microscopic realization of heterogeneity can be highly relevant
- we cannot average generating functions over the heterogeneity, because we cannot capture it in a probability distribution ...

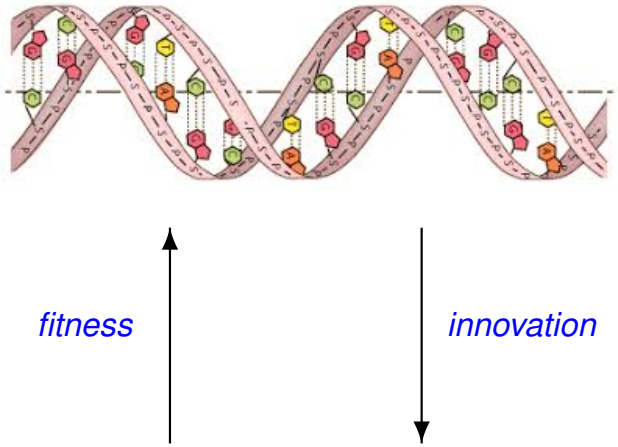
should we forget about disordered systems tools?
(replicas, generating functionals, cavity method, ...)

# How to capture heterogeneity in quantitative biology

*model how it is generated*

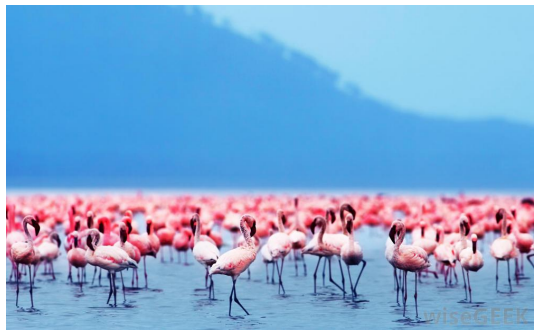
*slow dynamics:*

*evolving codes (genotypes* **y***)*




*fast dynamics:*

*interacting organisms (phenotypes* **x***)*



- coupled stochastic dynamics on *adiabatically separated timescales*
  - phenotype dynamics: genotype + population interactions + environment
  - genotype dynamics: phenotype fitness + hardware constraints

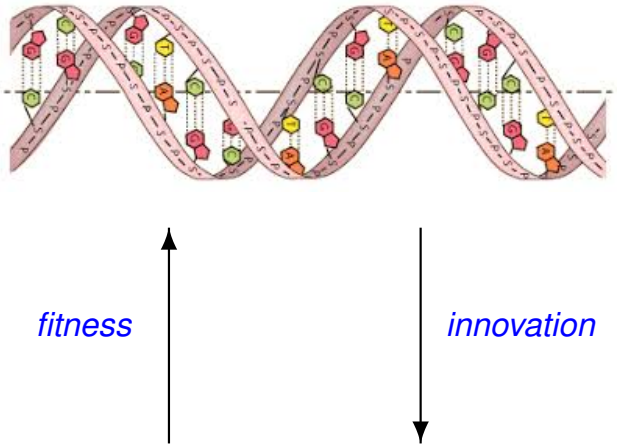
$$\textit{population}: \qquad P_t(\mathbf{x},\mathbf{y}) = \frac{1}{N_t}\sum_{i=1}^{N_t} \delta[\mathbf{x}-\mathbf{x}_i(t)]\;\delta[\mathbf{y}-\mathbf{y}_i(t)]$$

# How to capture heterogeneity in quantitative biology

*model how it is generated*

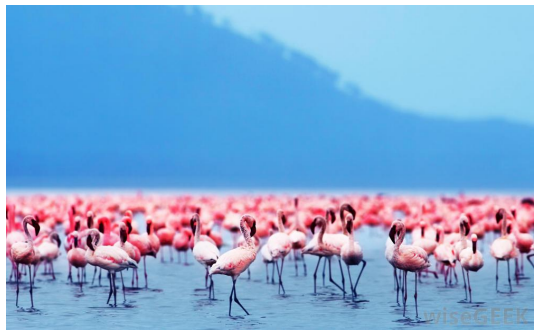
*slow dynamics:*

*evolving codes (genotypes* **y***)*




*fast dynamics:*

*interacting organisms (phenotypes* **x***)*



- coupled stochastic dynamics on *adiabatically separated timescales*
  - phenotype dynamics: genotype + population interactions + environment
  - genotype dynamics: phenotype fitness + hardware constraints

$$\textit{population}: \qquad P_t(\mathbf{x},\mathbf{y}) = \frac{1}{N_t}\sum_{i=1}^{N_t} \delta[\mathbf{x}-\mathbf{x}_i(t)]\;\delta[\mathbf{y}-\mathbf{y}_i(t)]$$

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-$n$ replica approach to biological evolution

4 Open questions

5 Summary

# Ising spin systems with slowly evolving interactions

(inspired by learning in neural networks)

- standard spin dynamics,
  evolving to equilibrium state

$$p(\boldsymbol{\sigma}|\mathbf{J}) = \frac{1}{Z(\mathbf{J})} \mathrm{e}^{-\beta H(\boldsymbol{\sigma},\mathbf{J})}, \qquad H(\boldsymbol{\sigma},\mathbf{J}) = -\sum_{i<j} \sigma_i J_{ij} \sigma_j - \sum_i \sigma_i \theta_i$$

$$Z(\mathbf{J}) = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma},\mathbf{J})}, \qquad F(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J})$$

- slow stochastic bond dynamics,
  Langevin eqn

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = \frac{1}{N} \langle \sigma_i \sigma_j \rangle_{\mathrm{spins}} - \mu J_{ij} + \frac{1}{\sqrt{N}} \eta_{ij}(t)$$

$$\langle \eta_{ij}(t) \rangle = 0, \quad \langle \eta_{ij}(t) \eta_{k\ell}(t') \rangle = 2\tau \tilde{T} \delta_{(i,j),(k,\ell)} \delta(t-t')$$

# Ising spin systems with slowly evolving interactions

(inspired by learning in neural networks)

- standard spin dynamics,
  evolving to equilibrium state

$$p(\boldsymbol{\sigma}|\mathbf{J}) = \frac{1}{Z(\mathbf{J})} \mathrm{e}^{-\beta H(\boldsymbol{\sigma},\mathbf{J})}, \qquad H(\boldsymbol{\sigma},\mathbf{J}) = -\sum_{i<j} \sigma_i J_{ij} \sigma_j - \sum_i \sigma_i \theta_i$$

$$Z(\mathbf{J}) = \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma},\mathbf{J})}, \qquad F(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J})$$

- slow stochastic bond dynamics,
  Langevin eqn

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = \frac{1}{N} \langle \sigma_i \sigma_j \rangle_{\mathrm{spins}} - \mu J_{ij} + \frac{1}{\sqrt{N}} \eta_{ij}(t)$$

$$\langle \eta_{ij}(t) \rangle = 0, \quad \langle \eta_{ij}(t) \eta_{k\ell}(t') \rangle = 2\tau \tilde{T} \delta_{(i,j),(k,\ell)} \delta(t-t')$$

# stationary state of coupled system?

- use

$$\langle \sigma_i \sigma_j \rangle_{\text{spins}} = -\frac{\partial}{\partial J_{ij}} F(\mathbf{J}) : \qquad N\tau \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{\partial}{\partial J_{ij}} \mathcal{H}(\mathbf{J}) + \sqrt{N}\eta_{ij}(t)$$

$$\mathcal{H}(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2$$

  Gibbs-Boltzmann state

$$\mathcal{P}_\infty(\mathbf{J}) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})}$$

- physics at largest timescales

$$\tilde{Z} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\left[-\frac{1}{\beta}\log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2\right]}$$

$$\tilde{F} = -\frac{1}{n\beta} \log \int \mathrm{d}\mathbf{J}\ Z^n(\mathbf{J}) \mathrm{e}^{-\frac{1}{2}\tilde{\beta}\mu N \sum_{i<j} J_{ij}^2}$$

  replica formula of disorder-averaged free energy
  of SK model (modulo a constant), but with $n = T/\tilde{T}$ ...
  (compute for integer $n$, analytical continuation to real $n$)

*analytically solvable!*

# stationary state of coupled system?

- use

$$\langle \sigma_i \sigma_j \rangle_{\text{spins}} = -\frac{\partial}{\partial J_{ij}} F(\mathbf{J}) : \qquad N\tau \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{\partial}{\partial J_{ij}} \mathcal{H}(\mathbf{J}) + \sqrt{N}\eta_{ij}(t)$$

$$\mathcal{H}(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2$$

  Gibbs-Boltzmann state

$$\mathcal{P}_\infty(\mathbf{J}) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})}$$

- physics at largest timescales

$$\tilde{Z} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\left[-\frac{1}{\beta}\log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2\right]}$$

$$\tilde{F} = -\frac{1}{n\beta} \log \int \mathrm{d}\mathbf{J}\ Z^n(\mathbf{J}) \mathrm{e}^{-\frac{1}{2}\tilde{\beta}\mu N \sum_{i<j} J_{ij}^2}$$

  replica formula of disorder-averaged free energy
  of SK model (modulo a constant), but with $n = T/\tilde{T}$ ...
  (compute for integer $n$, analytical continuation to real $n$)

*analytically solvable!*

# stationary state of coupled system?

- use

$$\langle \sigma_i \sigma_j \rangle_{\text{spins}} = -\frac{\partial}{\partial J_{ij}} F(\mathbf{J}) : \qquad N\tau \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{\partial}{\partial J_{ij}} \mathcal{H}(\mathbf{J}) + \sqrt{N}\eta_{ij}(t)$$

$$\mathcal{H}(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2$$

Gibbs-Boltzmann state

$$\mathcal{P}_\infty(\mathbf{J}) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})}$$

- physics at largest timescales

$$\tilde{Z} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\left[-\frac{1}{\beta}\log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2\right]}$$

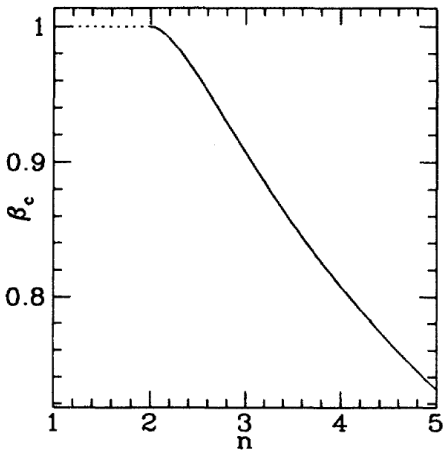
$$\tilde{F} = -\frac{1}{n\beta} \log \int \mathrm{d}\mathbf{J}\ Z^n(\mathbf{J}) \mathrm{e}^{-\frac{1}{2}\tilde{\beta}\mu N \sum_{i<j} J_{ij}^2}$$

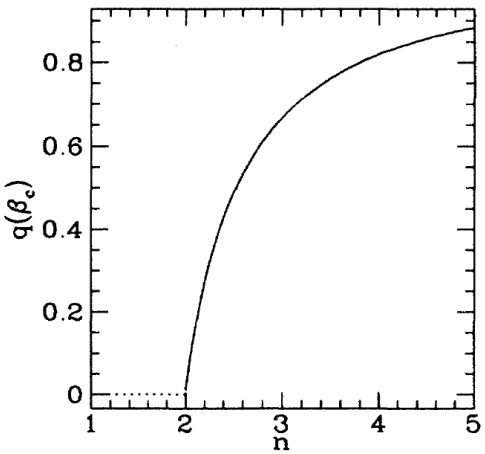
replica formula of disorder-averaged free energy
of SK model (modulo a constant), but with $n = T/\tilde{T}$ ...
(compute for integer $n$, analytical continuation to real $n$)

*analytically solvable!*

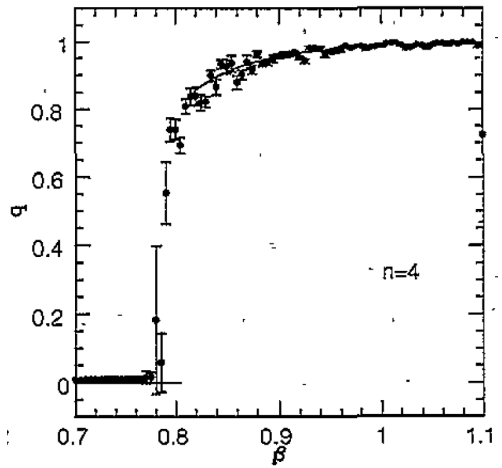
$n = T/\tilde{T}$ small: bond and spin dynamics weakly coupled
RSB, second order transitions

$n = T/\tilde{T}$ large: bond and spin dynamics strongly coupled
no RSB, first order transitions




P$\rightarrow$SG transition




first order for $n > 2$

*ACCC et al, Phys Rev B48, 1993*
*RW Penney et al, J Phys A26, 1993*

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-$n$ replica approach to biological evolution

4 Open questions

5 Summary

# XY spins and added layers of complexity

- standard spin dynamics,
  evolving to equilibrium state

$$p(\{\mathbf{S}\}|\mathbf{J}) = \frac{1}{Z(\mathbf{J})} \mathrm{e}^{-\beta H(\{\mathbf{S}\},\mathbf{J})}, \qquad H(\{\mathbf{S}\},\mathbf{J}) = -\sum_{i<j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$$

$$Z(\mathbf{J}) = \int \mathrm{d}\mathbf{S}_1 \ldots \mathrm{d}\mathbf{S}_N \, \mathrm{e}^{-\beta H(\{\mathbf{S}\},\mathbf{J})}, \qquad F(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J})$$

- slow stochastic bond dynamics,
  Langevin eqn

$$\frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = \frac{1}{N} \langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_{\mathrm{spins}} + \mu B_{ij} - \mu J_{ij} + \frac{1}{\sqrt{N}} \eta_{ij}(t)$$

$$\langle \eta_{ij}(t) \rangle = 0, \quad \langle \eta_{ij}(t) \eta_{k\ell}(t') \rangle = 2\tilde{T} \delta_{(i,j),(k,\ell)} \delta(t-t')$$

$$\textit{disorder}: \quad P(B_{ij}) = (2\pi \tilde{B}/N)^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}(B_{ij}-B_0/N)^2/(\tilde{B}/N)}$$

# XY spins and added layers of complexity

- standard spin dynamics,
  evolving to equilibrium state

$$p(\{\mathbf{S}\}|\mathbf{J}) = \frac{1}{Z(\mathbf{J})} \mathrm{e}^{-\beta H(\{\mathbf{S}\},\mathbf{J})}, \qquad H(\{\mathbf{S}\},\mathbf{J}) = -\sum_{i<j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j$$

$$Z(\mathbf{J}) = \int \mathrm{d}\mathbf{S}_1 \ldots \mathrm{d}\mathbf{S}_N \, \mathrm{e}^{-\beta H(\{\mathbf{S}\},\mathbf{J})}, \qquad F(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J})$$

- slow stochastic bond dynamics,
  Langevin eqn

$$\frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = \frac{1}{N} \langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_{\mathrm{spins}} + \mu B_{ij} - \mu J_{ij} + \frac{1}{\sqrt{N}} \eta_{ij}(t)$$

$$\langle \eta_{ij}(t) \rangle = 0, \quad \langle \eta_{ij}(t) \eta_{k\ell}(t') \rangle = 2\tilde{T} \delta_{(i,j),(k,\ell)} \delta(t - t')$$

$$\textit{disorder}: \quad P(B_{ij}) = (2\pi \tilde{B}/N)^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}(B_{ij} - B_0/N)^2/(\tilde{B}/N)}$$

# stationary state of coupled system?

- use

$$\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_{\text{spins}} = -\frac{\partial}{\partial J_{ij}} F(\mathbf{J}) : \qquad \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{1}{N} \frac{\partial}{\partial J_{ij}} \mathcal{H}(\mathbf{J}) + \frac{\eta_{ij}(t)}{\sqrt{N}}$$

$$\mathcal{H}(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2} \mu N \sum_{i<j} J_{ij}^2 - \mu N \sum_{i<j} B_{ij} J_{ij}$$

  Gibbs-Boltzmann state

$$\mathcal{P}_\infty(\mathbf{J}) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{\beta} \mathcal{H}(\mathbf{J})}$$

- physics at largest timescales

$$\tilde{Z} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta} \mathcal{H}(\mathbf{J})} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta} \left[ -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2} \mu N \sum_{i<j} J_{ij}^2 - \mu N \sum_{i<j} B_{ij} J_{ij} \right]}$$

$$\tilde{F} = -\frac{1}{n\beta} \log \int \mathrm{d}\mathbf{J}\ Z^n(\mathbf{J}) \mathrm{e}^{\mu \tilde{\beta} N \sum_{i<j} B_{ij} J_{ij} - \frac{1}{2} \tilde{\beta} \mu N \sum_{i<j} J_{ij}^2}$$

  replica formula of disorder-averaged free energy
  of XY-spin model (modulo a constant), with $n = T/\tilde{T}$ ...
  (compute for integer $n$, analytical continuation to real $n$)

*analytically solvable!*

# stationary state of coupled system?

- use

$$\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_{\text{spins}} = -\frac{\partial}{\partial J_{ij}} F(\mathbf{J}) : \qquad \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{1}{N} \frac{\partial}{\partial J_{ij}} \mathcal{H}(\mathbf{J}) + \frac{\eta_{ij}(t)}{\sqrt{N}}$$

$$\mathcal{H}(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2} \mu N \sum_{i<j} J_{ij}^2 - \mu N \sum_{i<j} B_{ij} J_{ij}$$

Gibbs-Boltzmann state

$$\mathcal{P}_\infty(\mathbf{J}) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{\beta} \mathcal{H}(\mathbf{J})}$$

- physics at largest timescales

$$\tilde{Z} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\mathcal{H}(\mathbf{J})} = \int \mathrm{d}\mathbf{J}\ \mathrm{e}^{-\tilde{\beta}\left[-\frac{1}{\beta}\log Z(\mathbf{J}) + \frac{1}{2}\mu N \sum_{i<j} J_{ij}^2 - \mu N \sum_{i<j} B_{ij} J_{ij}\right]}$$

$$\tilde{F} = -\frac{1}{n\beta} \log \int \mathrm{d}\mathbf{J}\ Z^n(\mathbf{J}) \mathrm{e}^{\mu\tilde{\beta} N \sum_{i<j} B_{ij} J_{ij} - \frac{1}{2}\tilde{\beta}\mu N \sum_{i<j} J_{ij}^2}$$

replica formula of disorder-averaged free energy
of XY-spin model (modulo a constant), with $n = T/\tilde{T}$ ...
(compute for integer $n$, analytical continuation to real $n$)

*analytically solvable!*

# stationary state of coupled system?

- use

$$\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle_{\text{spins}} = -\frac{\partial}{\partial J_{ij}} F(\mathbf{J}) : \qquad \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{1}{N} \frac{\partial}{\partial J_{ij}} \mathcal{H}(\mathbf{J}) + \frac{\eta_{ij}(t)}{\sqrt{N}}$$

$$\mathcal{H}(\mathbf{J}) = -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2} \mu N \sum_{i<j} J_{ij}^2 - \mu N \sum_{i<j} B_{ij} J_{ij}$$

  Gibbs-Boltzmann state

$$\mathcal{P}_\infty(\mathbf{J}) = \tilde{Z}^{-1} \mathrm{e}^{-\tilde{\beta} \mathcal{H}(\mathbf{J})}$$

- physics at largest timescales

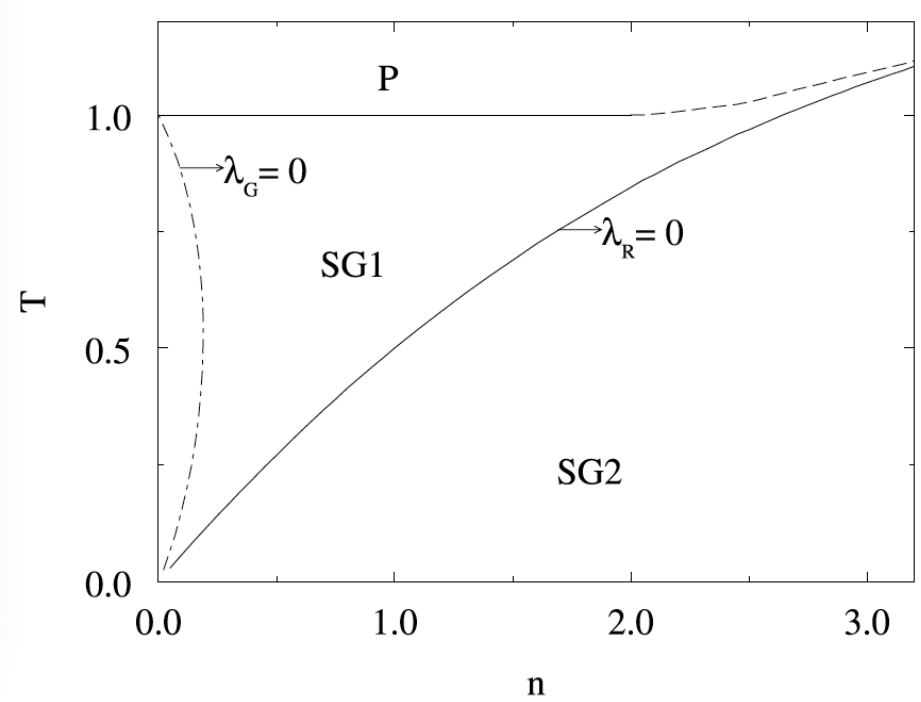
$$\tilde{Z} = \int \mathrm{d}\mathbf{J}\, \mathrm{e}^{-\tilde{\beta} \mathcal{H}(\mathbf{J})} = \int \mathrm{d}\mathbf{J}\, \mathrm{e}^{-\tilde{\beta} \left[ -\frac{1}{\beta} \log Z(\mathbf{J}) + \frac{1}{2} \mu N \sum_{i<j} J_{ij}^2 - \mu N \sum_{i<j} B_{ij} J_{ij} \right]}$$

$$\tilde{F} = -\frac{1}{n\beta} \log \int \mathrm{d}\mathbf{J}\, Z^n(\mathbf{J}) \mathrm{e}^{\mu \tilde{\beta} N \sum_{i<j} B_{ij} J_{ij} - \frac{1}{2} \tilde{\beta} \mu N \sum_{i<j} J_{ij}^2}$$

  replica formula of disorder-averaged free energy
  of XY-spin model (modulo a constant), with $n = T/\tilde{T}$ ...
  (compute for integer $n$, analytical continuation to real $n$)

*analytically solvable!*

$n = T/\tilde{T}$




two SG phases

SG1: freezing on spin timescales only, $q_1 > 0$, $q_0 = 0$
SG2: freezing on all timescales, $q_1 > 0$, $q_0 > 0$

$$q_0 = \lim_{N\to\infty} \frac{1}{N} \sum_i \left\langle \overline{\langle \mathbf{S}_i \rangle}^2 \right\rangle_B, \qquad q_1 = \lim_{N\to\infty} \frac{1}{N} \sum_i \left\langle \overline{\langle \mathbf{S}_i \rangle^2} \right\rangle_B$$

*G Jongen et al, J Phys A31, 1998*
*G Jongen et al, J Phys A34, 2001*

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-*n* replica approach to biological evolution

4 Open questions

5 Summary

# Link with the Parisi scheme

SK model:

$$H = -\sum_{i<j} J_{ij}\sigma_i\sigma_j, \quad \textit{random bonds}: \quad \overline{J_{ij}} = \frac{J_0}{N}, \quad \overline{J_{ij}^2} - \overline{J_{ij}}^2 = \frac{J^2}{N}$$

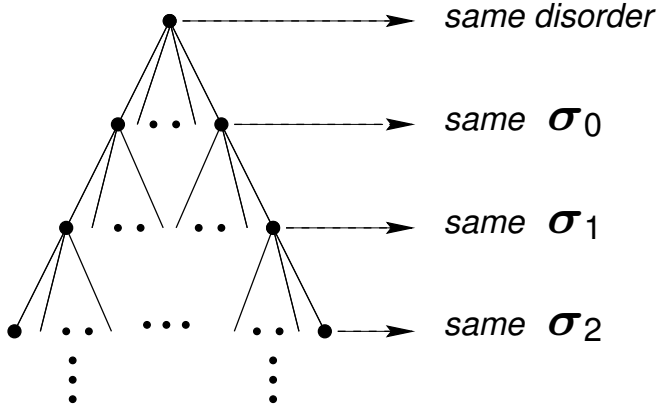
*intuition:* *'slow' spins act as slowly evolving interactions between 'fast' spins ...*

- assume spins evolving on disparate timescales

  $\{1,\ldots,N\} = \bigcup_{\ell=0}^{L} I_\ell$

  $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_0,\ldots,\boldsymbol{\sigma}_L)$

  $\boldsymbol{\sigma}_\ell = \{\sigma_j |\ j \in I_\ell\}, \quad \tau_\ell \ll \tau_{\ell-1}$




- at each level $\ell$:
  minimize free energy (Boltzmann's H-theorem)

  – Boltzmann form at each level $\ell$
  – Hamiltonian at level $\ell$: free energy of level $\ell+1$
  – constrained entropy at each level $\ell$: distinct temperatures $T_\ell$

# Link with the Parisi scheme

SK model:

$$H = -\sum_{i<j} J_{ij}\sigma_i\sigma_j, \quad \textit{random bonds}: \quad \overline{J_{ij}} = \frac{J_0}{N}, \quad \overline{J_{ij}^2} - \overline{J_{ij}}^2 = \frac{J^2}{N}$$

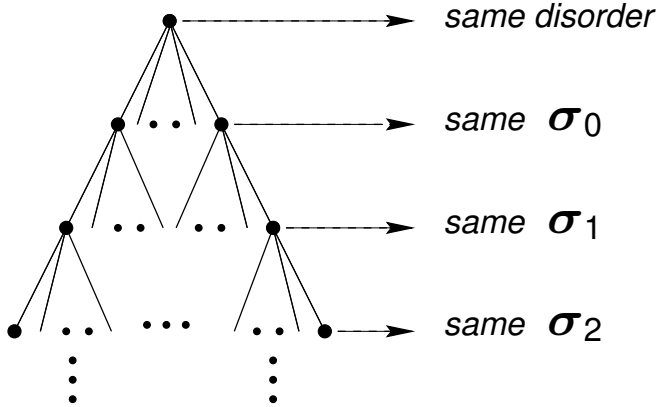
*intuition:* *'slow' spins act as slowly evolving interactions between 'fast' spins ...*

- assume spins evolving on disparate timescales

  $\{1,\ldots,N\} = \bigcup_{\ell=0}^{L} I_\ell$

  $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_0,\ldots,\boldsymbol{\sigma}_L)$

  $\boldsymbol{\sigma}_\ell = \{\sigma_j |\ j \in I_\ell\}, \quad \tau_\ell \ll \tau_{\ell-1}$




- at each level $\ell$:
  minimize free energy (Boltzmann's H-theorem)

  – Boltzmann form at each level $\ell$
  – Hamiltonian at level $\ell$: free energy of level $\ell+1$
  – constrained entropy at each level $\ell$: distinct temperatures $T_\ell$

- resulting theory

$$\textit{level } L,\ \textit{fastest spins}: \qquad Z_L = \mathrm{Tr}_{\boldsymbol{\sigma}_L} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})},$$

$$\textit{higher levels } \ell < L: \qquad Z_\ell = \mathrm{Tr}_{\boldsymbol{\sigma}_\ell} [Z_{\ell+1}]^{\tilde{m}_{\ell+1}}, \qquad \tilde{m}_\ell = \frac{\beta_{\ell-1}}{\beta_\ell}, \quad \beta_L = \beta$$

$\tilde{m}_\ell$ : follow from values of constrained entropies $S_\ell$

- disorder-averaged free energy, at largest timescale

$$\overline{F_0} = -\frac{1}{\beta_0} \overline{\log Z_0} = -\lim_{\tilde{n}\to 0} \frac{1}{\tilde{n}\beta_0} \log \overline{Z_0^{\tilde{n}}}$$

- assume ergodicity at each level $\ell$, extremize $f = \lim_{N\to\infty}(\overline{F_0}/N)$ over $\epsilon_\ell = |I_\ell|/N$:

$$\epsilon_L \to 1, \quad \epsilon_{\ell<L} \to 0: \qquad \textit{slow spins are vanishing fraction, entropy densities } S_\ell/N \to 0$$

$$\textit{for } L\to\infty: \quad \textit{Parisi's full RSB scheme}, \quad \textit{with } m_\ell = \prod_{k=\ell}^{L} \tilde{m}_k,$$

*J Van Mourik and ACCC, J Phys A34, 2001*

- resulting theory

$$\textit{level } L,\ \textit{fastest spins}: \qquad Z_L = \mathrm{Tr}_{\boldsymbol{\sigma}_L} \mathrm{e}^{-\beta H(\boldsymbol{\sigma})},$$

$$\textit{higher levels } \ell < L: \qquad Z_\ell = \mathrm{Tr}_{\boldsymbol{\sigma}_\ell} [Z_{\ell+1}]^{\tilde{m}_{\ell+1}}, \qquad \tilde{m}_\ell = \frac{\beta_{\ell-1}}{\beta_\ell}, \quad \beta_L = \beta$$

$\tilde{m}_\ell$ : follow from values of constrained entropies $S_\ell$

- disorder-averaged free energy, at largest timescale

$$\overline{F_0} = -\frac{1}{\beta_0} \overline{\log Z_0} = -\lim_{\tilde{n}\to 0} \frac{1}{\tilde{n}\beta_0} \log \overline{Z_0^{\tilde{n}}}$$

- assume ergodicity at each level $\ell$, extremize $f = \lim_{N\to\infty}(\overline{F_0}/N)$ over $\epsilon_\ell = |I_\ell|/N$:

$$\epsilon_L \to 1, \quad \epsilon_{\ell<L} \to 0: \qquad \textit{slow spins are vanishing fraction, entropy densities } S_\ell/N \to 0$$

$$\textit{for } L\to\infty: \quad \textit{Parisi's full RSB scheme}, \quad \textit{with } m_\ell = \prod_{k=\ell}^{L} \tilde{m}_k,$$

*J Van Mourik and ACCC, J Phys A34, 2001*

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-*n* replica approach to biological evolution

4 Open questions

5 Summary

# Self-programming in neural systems

bonds=programme evolving at multiple timescales

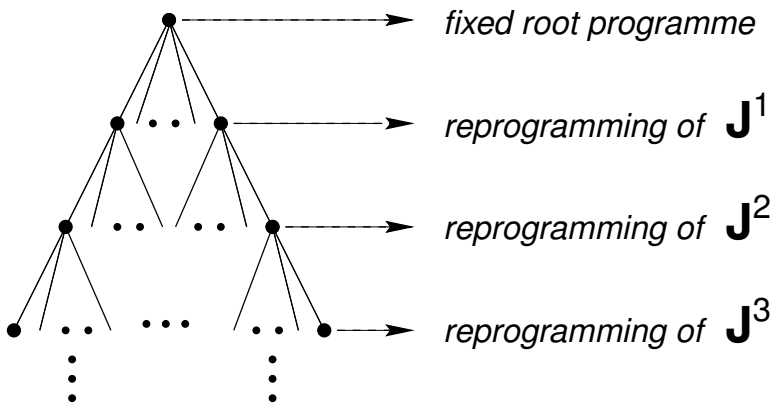
- binary neurons $\boldsymbol{\sigma} \in \{-1,1\}^N$,
  evolving to equilibrium state

$$p(\boldsymbol{\sigma}|\mathbf{J}) = \frac{1}{Z(\mathbf{J})} \mathrm{e}^{-\beta H(\boldsymbol{\sigma},\mathbf{J})}, \qquad H(\boldsymbol{\sigma},\mathbf{J}) = -\sum_{i<j} \sigma_i J_{ij} \sigma_j + \frac{1}{2} N \sum_{i<j} \mu_{ij} J_{ij}^2$$

- slow stochastic bond dynamics
  $\ell = 1 \ldots L$ reprogramming levels

  $I_\ell = \{(i,j) |\ \tau_{ij} = \tau_\ell,\ T_{ij} = T_\ell,\ \mu_{ij} = \mu_\ell\}$

  $\mathbf{J}^\ell = \{J_{ij} |\ (i,j) \in I_\ell$




$$(i,j) \in I_\ell : \qquad \tau_\ell \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = \frac{1}{N} \langle \sigma_i \sigma_j \rangle - \mu_\ell J_{ij} + \frac{\sqrt{\tau_\ell}}{\sqrt{N}} \eta_{ij}(t)$$

$$\langle \eta_{ij}(t) \rangle = 0, \quad \langle \eta_{ij}(t) \eta_{k\ell}(t') \rangle = 2 T_\ell \delta_{(i,j),(k,\ell)} \delta(t - t')$$

# stationary state of coupled system?

- at level $\ell$:

$$\tau_\ell \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{1}{N}\frac{\partial}{\partial J_{ij}} H_\ell(\mathbf{J}^\ell, \ldots, \mathbf{J}^L) + \sqrt{\frac{\tau_\ell}{N}}$$

$$H_1(\ldots) = -\frac{1}{\beta}\log \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma}, \mathbf{J})}, \qquad H_{\ell+1}(\ldots) = -\frac{1}{\beta_\ell}\log Z_\ell[\ldots]$$

$$Z_\ell[\mathbf{J}^{\ell+1}, \ldots, \mathbf{J}^L] = \int \mathrm{d}\mathbf{J}^\ell \; \mathrm{e}^{-\beta_\ell H_\ell(\mathbf{J}^\ell, \ldots, \mathbf{J}^L)}$$

- disorder:
  level membership of bonds

$$\epsilon_{ij}(\ell) \in \{0, 1\}: \qquad \text{i.i.d.r.v}, \;\; \mathrm{Prob}[\epsilon_{ij}(\ell) = 1] = \epsilon_\ell, \qquad \sum_{\ell=1}^{L} \epsilon_\ell = 1$$

- physics at largest timescales

$$\mathcal{F} = -\frac{1}{\beta_L}\overline{\log Z_L} = -\lim_{m_{L+1}\to 0} \frac{1}{m_{L+1}\beta_L} \log \overline{Z_L^{m_{L+1}}}$$

  (compute for integer $m_{L+1}$, analytical continuation to real $m_{L+1}$)

analytically solvable!

## stationary state of coupled system?

- at level $\ell$:

$$\tau_\ell \frac{\mathrm{d}}{\mathrm{d}t} J_{ij} = -\frac{1}{N} \frac{\partial}{\partial J_{ij}} H_\ell(\mathbf{J}^\ell, \ldots, \mathbf{J}^L) + \sqrt{\frac{\tau_\ell}{N}}$$

$$H_1(\ldots) = -\frac{1}{\beta} \log \sum_{\boldsymbol{\sigma}} \mathrm{e}^{-\beta H(\boldsymbol{\sigma}, \mathbf{J})}, \qquad H_{\ell+1}(\ldots) = -\frac{1}{\beta_\ell} \log Z_\ell[\ldots]$$

$$Z_\ell[\mathbf{J}^{\ell+1}, \ldots, \mathbf{J}^L] = \int \mathrm{d}\mathbf{J}^\ell \; \mathrm{e}^{-\beta_\ell H_\ell(\mathbf{J}^\ell, \ldots, \mathbf{J}^L)}$$

- disorder:
  level membership of bonds

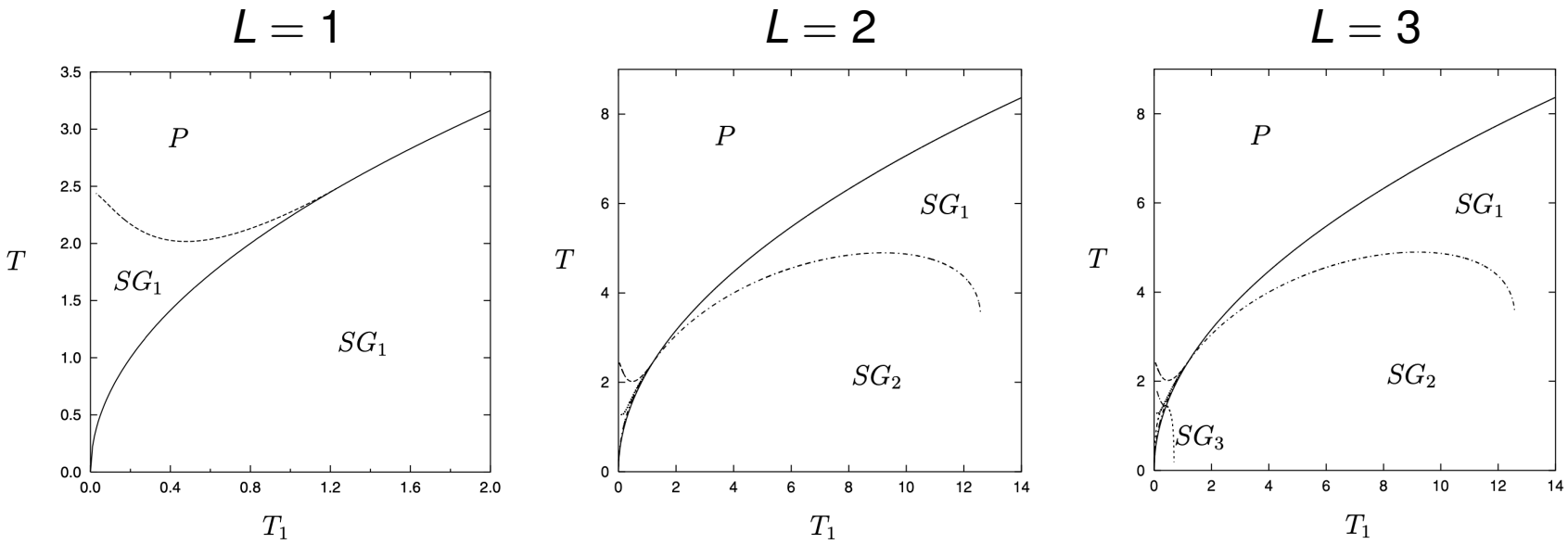
$$\epsilon_{ij}(\ell) \in \{0, 1\} : \qquad \text{i.i.d.r.v}, \;\; \mathrm{Prob}[\epsilon_{ij}(\ell) = 1] = \epsilon_\ell, \qquad \sum_{\ell=1}^{L} \epsilon_\ell = 1$$

- physics at largest timescales

$$\mathcal{F} = -\frac{1}{\beta_L} \overline{\log Z_L} = -\lim_{m_{L+1} \to 0} \frac{1}{m_{L+1} \beta_L} \log \overline{Z_L^{m_{L+1}}}$$

  (compute for integer $m_{L+1}$, analytical continuation to real $m_{L+1}$)

*analytically solvable!*




- increasingly complex phase diagrams
- multiple SG states, indicating freezing on distinct timescales
- first order transitions
- re-entrance phenomena

*T Uezu and ACCC, J Phys A35, 2002*

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-*n* replica approach to biological evolution

4 Open questions

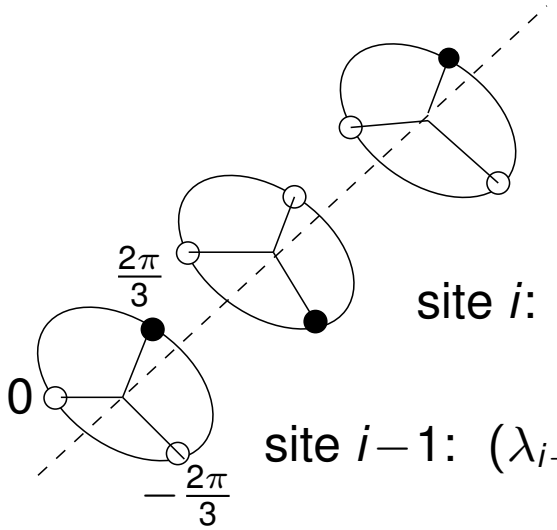
5 Summary

# Biological evolution: genesis of protein structure

simple model of heteropolymer with slowly evolving composition

genotype: $(\lambda_1, \ldots, \lambda_N)$ (primary structure)

phenotype: $(\phi_1, \ldots, \phi_N)$ (secondary structure)




$$H(\phi|\lambda) = -\frac{J_p}{N}\sum_{ij}\overbrace{\xi(\lambda_i)\xi(\lambda_j)\,\delta_{\phi_i,\phi_j}}^{\text{polarity forces}} - J_s\sum_i \underbrace{\cos[(\phi_{i+1}-\phi_i)-(\phi_i-\phi_{i-1})-a(\lambda_i)]}_{\text{steric forces}}$$

$\lambda_i \in \{1, \ldots, 20\}$: amino-acid at site $i$

$\xi(\lambda) \in \mathbb{R}$: polarity of amino-acid $\lambda$

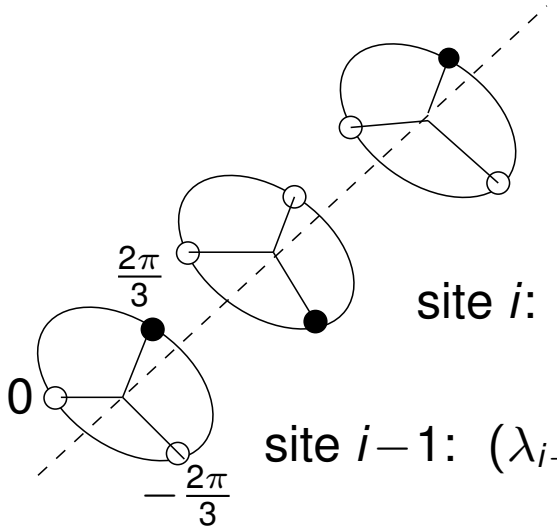
$a(\lambda) \in [-\pi, \pi]$: winding angle of amino-acid $\lambda$

# Biological evolution: genesis of protein structure

simple model of heteropolymer with slowly evolving composition

genotype: $(\lambda_1, \ldots, \lambda_N)$
(primary structure)

phenotype: $(\phi_1, \ldots, \phi_N)$
(secondary structure)




$$H(\phi|\lambda) = \overbrace{-\frac{J_p}{N}\sum_{ij}\xi(\lambda_i)\xi(\lambda_j)\,\delta_{\phi_i,\phi_j}}^{\textit{polarity forces}} \underbrace{-J_s\sum_i \cos[(\phi_{i+1}-\phi_i)-(\phi_i-\phi_{i-1})-a(\lambda_i)]}_{\textit{steric forces}}$$

$\lambda_i \in \{1, \ldots, 20\}$: amino-acid at site $i$

$\xi(\lambda) \in \mathbb{R}$: polarity of amino-acid $\lambda$
$a(\lambda) \in [-\pi, \pi]$: winding angle of amino-acid $\lambda$

slow dynamics of genotype (primary structure $\boldsymbol{\lambda}$):

- if $\boldsymbol{\lambda}$ real:

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_i = \Big\langle -\frac{\partial}{\partial\lambda_i}\Big[H(\boldsymbol{\phi}|\boldsymbol{\lambda}) + U(\boldsymbol{\lambda})\Big]\Big\rangle_{\boldsymbol{\phi}} + \eta_i(t)$$

$$\langle\eta_i(t)\rangle = 0, \quad \langle\eta_i(t)\eta_j(t')\rangle = 2\tilde{T}\delta_{ij}\delta(t-t')$$

$U(\boldsymbol{\lambda})$ : *utility potential*

*rationale:*

*minimize* $H(\boldsymbol{\phi}|\boldsymbol{\lambda})$ $\rightarrow$ *protein folds*
*minimize* $U(\boldsymbol{\lambda})$ $\rightarrow$ *protein is useful*

- use

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_i = -\frac{\partial}{\partial\lambda_i}H_{\mathrm{eff}}(\boldsymbol{\lambda}) + \eta_i(t)$$

$$H_{\mathrm{eff}}(\boldsymbol{\lambda}) = U(\boldsymbol{\lambda}) - \beta^{-1}\log Z_{\mathrm{f}}(\boldsymbol{\lambda}), \qquad Z_f(\boldsymbol{\lambda}) = \sum_{\boldsymbol{\phi}} \mathrm{e}^{-\beta H(\boldsymbol{\phi}|\boldsymbol{\lambda})}$$

- physics at genetic timescale

$$f_N = -\frac{1}{\tilde{\beta}N}\log\sum_{\boldsymbol{\lambda}} \mathrm{e}^{-\tilde{\beta}H_{\mathrm{eff}}(\boldsymbol{\lambda})} = -\frac{1}{n\beta N}\log\sum_{\boldsymbol{\lambda}} Z_{\mathrm{f}}^n(\boldsymbol{\lambda})\mathrm{e}^{-n\beta U(\boldsymbol{\lambda})}$$

$$= -\frac{1}{n\beta N}\log\sum_{\boldsymbol{\phi}^1\ldots\boldsymbol{\phi}^n}\sum_{\boldsymbol{\lambda}} \mathrm{e}^{-\beta\sum_{\alpha=1}^n H(\boldsymbol{\phi}^\alpha|\boldsymbol{\lambda}) - n\beta U(\boldsymbol{\lambda})}$$

slow dynamics of genotype (primary structure $\boldsymbol{\lambda}$):

- if $\boldsymbol{\lambda}$ real:

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_i = \Big\langle -\frac{\partial}{\partial\lambda_i}\Big[H(\boldsymbol{\phi}|\boldsymbol{\lambda}) + U(\boldsymbol{\lambda})\Big]\Big\rangle_{\boldsymbol{\phi}} + \eta_i(t)$$

$$\langle\eta_i(t)\rangle = 0, \quad \langle\eta_i(t)\eta_j(t')\rangle = 2\tilde{T}\delta_{ij}\delta(t-t')$$

$U(\boldsymbol{\lambda})$ : *utility potential*

*rationale:*

*minimize* $H(\boldsymbol{\phi}|\boldsymbol{\lambda})$ $\rightarrow$ *protein folds*
*minimize* $U(\boldsymbol{\lambda})$ $\rightarrow$ *protein is useful*

- use

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_i = -\frac{\partial}{\partial\lambda_i}H_{\mathrm{eff}}(\boldsymbol{\lambda}) + \eta_i(t)$$

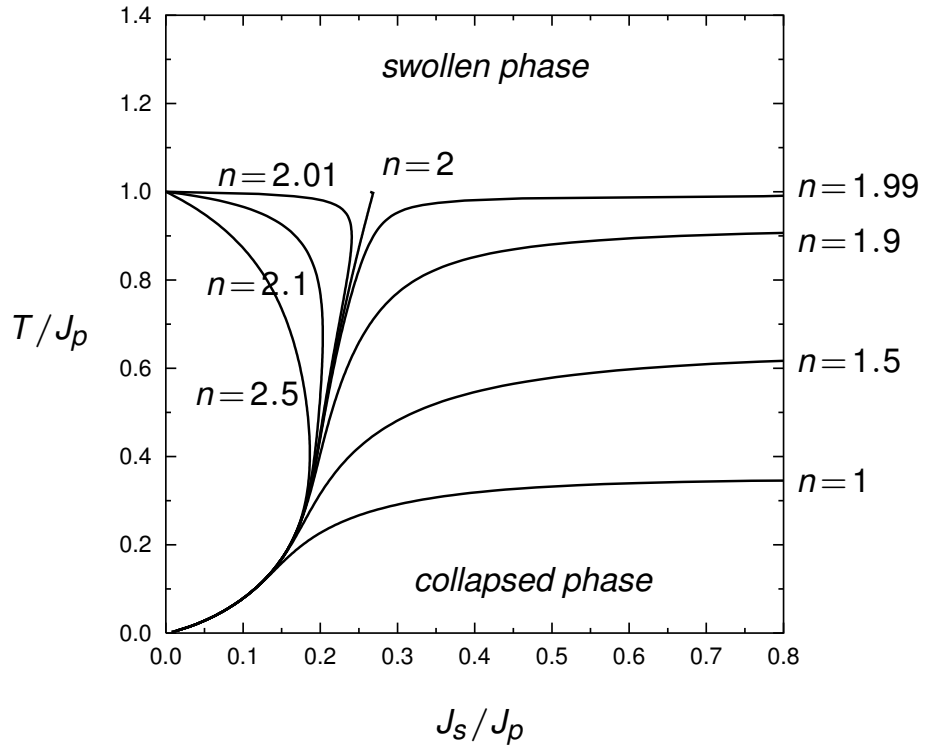
$$H_{\mathrm{eff}}(\boldsymbol{\lambda}) = U(\boldsymbol{\lambda}) - \beta^{-1}\log Z_{\mathrm{f}}(\boldsymbol{\lambda}), \qquad Z_f(\boldsymbol{\lambda}) = \sum_{\boldsymbol{\phi}} \mathrm{e}^{-\beta H(\boldsymbol{\phi}|\boldsymbol{\lambda})}$$

- physics at genetic timescale

$$f_N = -\frac{1}{\tilde{\beta}N}\log\sum_{\boldsymbol{\lambda}}\mathrm{e}^{-\tilde{\beta}H_{\mathrm{eff}}(\boldsymbol{\lambda})} = -\frac{1}{n\beta N}\log\sum_{\boldsymbol{\lambda}} z_{\mathrm{f}}^n(\boldsymbol{\lambda})\mathrm{e}^{-n\beta U(\boldsymbol{\lambda})}$$

$$= -\frac{1}{n\beta N}\log\sum_{\boldsymbol{\phi}^1\ldots\boldsymbol{\phi}^n}\sum_{\boldsymbol{\lambda}}\mathrm{e}^{-\beta\sum_{\alpha=1}^n H(\boldsymbol{\phi}^\alpha|\boldsymbol{\lambda}) - n\beta U(\boldsymbol{\lambda})}$$

calculations involve diagonalization of replicated transfer matrices and techniques from random field spin chains




Continuous bifurcations from swollen to collapsed states

$n > 2$: 1st order transition

*H Chakravorty et al, J Phys A35, 2002*
*S Rabello et al, J Phys A41, 2008*

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-*n* replica approach to biological evolution

4 Open questions

5 Summary

# Finite-$n$ replica approach to biological evolution

generalization of the previous problem ...

$$\begin{aligned} \textit{codes (genotypes)}: &\quad \mathbf{y}_1, \ldots, \mathbf{y}_N \\ \textit{organisms (phenotypes)}: &\quad \mathbf{x}_1, \ldots, \mathbf{x}_N \\ \textit{population}: &\quad P_t(\mathbf{x}, \mathbf{y}) = \frac{1}{N_t} \sum_{i=1}^{N_t} \delta[\mathbf{x} - \mathbf{x}_i(t)]\, \delta[\mathbf{y} - \mathbf{y}_i(t)] \end{aligned}$$

- fast phenotype dynamics:

$$\tau_x \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = -\nabla_x \Big[ \overbrace{H(\mathbf{x},\mathbf{y})}^{\text{genomic instructions}} + \overbrace{H(\mathbf{x}|P)}^{\text{environment}} \Big] + \overbrace{\eta(t)}^{\text{randomness}}$$

- slow genotype dynamics:
  - stochastic, time scales $\tau_y \gg \tau_x$
  - favour codes $\mathbf{y}$ that give 'fit' phenotypes
  - 'fit' $\mathbf{x}$: low values of $H(\mathbf{x},\mathbf{y}) + H(\mathbf{x}|P)$

# Finite-$n$ replica approach to biological evolution

generalization of the previous problem ...

$$\begin{aligned} \textit{codes (genotypes)}: &\quad \mathbf{y}_1, \ldots, \mathbf{y}_N \\ \textit{organisms (phenotypes)}: &\quad \mathbf{x}_1, \ldots, \mathbf{x}_N \\ \textit{population}: &\quad P_t(\mathbf{x}, \mathbf{y}) = \frac{1}{N_t} \sum_{i=1}^{N_t} \delta[\mathbf{x} - \mathbf{x}_i(t)]\, \delta[\mathbf{y} - \mathbf{y}_i(t)] \end{aligned}$$

- fast phenotype dynamics:

$$\tau_x \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = -\nabla_x \Big[ \overbrace{H(\mathbf{x}, \mathbf{y})}^{\textit{genomic instructions}} + \overbrace{H(\mathbf{x}|P)}^{\textit{environment}} \Big] + \overbrace{\eta(t)}^{\textit{randomness}}$$

- slow genotype dynamics:
  - stochastic, time scales $\tau_y \gg \tau_x$
  - favour codes $\mathbf{y}$ that give 'fit' phenotypes
  - 'fit' $\mathbf{x}$: low values of $H(\mathbf{x}, \mathbf{y}) + H(\mathbf{x}|P)$

# Finite-$n$ replica approach to biological evolution

generalization of the previous problem ...

$$\begin{aligned} \textit{codes (genotypes)}: &\quad \mathbf{y}_1, \ldots, \mathbf{y}_N \\ \textit{organisms (phenotypes)}: &\quad \mathbf{x}_1, \ldots, \mathbf{x}_N \\ \textit{population}: &\quad P_t(\mathbf{x}, \mathbf{y}) = \frac{1}{N_t} \sum_{i=1}^{N_t} \delta[\mathbf{x} - \mathbf{x}_i(t)]\, \delta[\mathbf{y} - \mathbf{y}_i(t)] \end{aligned}$$

- fast phenotype dynamics:

$$\tau_x \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = -\nabla_x \Big[ \overbrace{H(\mathbf{x}, \mathbf{y})}^{\textit{genomic instructions}} + \overbrace{H(\mathbf{x}|P)}^{\textit{environment}} \Big] + \overbrace{\eta(t)}^{\textit{randomness}}$$

- slow genotype dynamics:
  - stochastic, time scales $\tau_y \gg \tau_x$
  - favour codes $\mathbf{y}$ that give 'fit' phenotypes
  - 'fit' $\mathbf{x}$: low values of $H(\mathbf{x}, \mathbf{y}) + H(\mathbf{x}|P)$

## Resulting model

*genotype* **y** (*slow dynamics*) $\rightleftharpoons$ *phenotype* **x** (*fast dynamics*)

$$\tau_x \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = -\nabla_x \Big[ \overbrace{H(\mathbf{x},\mathbf{y})}^{\text{genomic instructions}} + \overbrace{H_x(\mathbf{x})}^{\text{environment}} \Big] + \overbrace{\eta_x(t)}^{\text{randomness}}$$

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \Big[ \underbrace{H(\mathbf{x},\mathbf{y})}_{\text{genomic instructions}} + \underbrace{H_y(\mathbf{y})}_{\text{coding hardware}} \Big] + \underbrace{\eta_y(t)}_{\text{mutations, crossover, } \ldots}$$

*solution?*

On evolutionary time scales:
(adiabatic separation: $\tau_x \ll \tau_y$)

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\Big\langle \nabla_y H(\mathbf{x},\mathbf{y}) \Big\rangle_{\mathbf{x}|\mathbf{y}} - \nabla_y H_y(\mathbf{y}) + \eta_y(t)$$

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z_x(\mathbf{y})} \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y}) + H_x(\mathbf{x})\right]}, \qquad Z_x(\mathbf{y}) = \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y}) + H_x(\mathbf{x})\right]}$$

## Resulting model

*genotype* **y** (*slow dynamics*) $\rightleftharpoons$ *phenotype* **x** (*fast dynamics*)

$$\tau_x \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = -\nabla_x \Big[ \overbrace{H(\mathbf{x},\mathbf{y})}^{\text{genomic instructions}} + \overbrace{H_x(\mathbf{x})}^{\text{environment}} \Big] + \overbrace{\eta_x(t)}^{\text{randomness}}$$

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \Big[ \underbrace{H(\mathbf{x},\mathbf{y})}_{\text{genomic instructions}} + \underbrace{H_y(\mathbf{y})}_{\text{coding hardware}} \Big] + \underbrace{\eta_y(t)}_{\text{mutations, crossover, ...}}$$

*solution?*

On evolutionary time scales:
(adiabatic separation: $\tau_x \ll \tau_y$)

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\Big\langle \nabla_y H(\mathbf{x},\mathbf{y}) \Big\rangle_{\mathbf{x}|\mathbf{y}} - \nabla_y H_y(\mathbf{y}) + \eta_y(t)$$

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z_x(\mathbf{y})} \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y}) + H_x(\mathbf{x})\right]}, \qquad Z_x(\mathbf{y}) = \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y}) + H_x(\mathbf{x})\right]}$$

## Resulting model

*genotype* **y** (*slow dynamics*) $\rightleftharpoons$ *phenotype* **x** (*fast dynamics*)

$$\tau_x \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = -\nabla_x \Big[ \overbrace{H(\mathbf{x},\mathbf{y})}^{\text{genomic instructions}} + \overbrace{H_x(\mathbf{x})}^{\text{environment}} \Big] + \overbrace{\eta_x(t)}^{\text{randomness}}$$

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \Big[ \underbrace{H(\mathbf{x},\mathbf{y})}_{\text{genomic instructions}} + \underbrace{H_y(\mathbf{y})}_{\text{coding hardware}} \Big] + \underbrace{\eta_y(t)}_{\text{mutations, crossover, } \ldots}$$

*solution?*

On evolutionary time scales:
(adiabatic separation: $\tau_x \ll \tau_y$)

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\Big\langle \nabla_y H(\mathbf{x},\mathbf{y}) \Big\rangle_{\mathbf{x}|\mathbf{y}} - \nabla_y H_y(\mathbf{y}) + \eta_y(t)$$

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z_x(\mathbf{y})} \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y}) + H_x(\mathbf{x})\right]}, \qquad Z_x(\mathbf{y}) = \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y}) + H_x(\mathbf{x})\right]}$$

$$\left\langle \nabla_y H(\mathbf{x},\mathbf{y}) \right\rangle_{\mathbf{x}|\mathbf{y}} = \frac{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} \nabla_y H(\mathbf{x},\mathbf{y})}{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}}$$

$$= -\frac{1}{\beta_x} \nabla_y \log \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} = \nabla_y \underbrace{F_x(\mathbf{y})}_{\textit{free energy}}$$

hence

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \left[F_x(\mathbf{y}) + H_y(\mathbf{y})\right] + \eta_y(t)$$

$$P_\infty(\mathbf{y}) = \frac{1}{Z_y} \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

Equilibrium:
*genotypic free energy*

$$F_y = -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

$$= -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y H_y(\mathbf{y})} \left[Z_x(\mathbf{y})\right]^{\frac{\beta_y}{\beta_x}}$$

$$\left\langle \nabla_y H(\mathbf{x},\mathbf{y}) \right\rangle_{\mathbf{x}|\mathbf{y}} = \frac{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} \nabla_y H(\mathbf{x},\mathbf{y})}{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}}$$

$$= -\frac{1}{\beta_x} \nabla_y \log \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} = \nabla_y \underbrace{F_x(\mathbf{y})}_{\textit{free energy}}$$

hence

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \left[F_x(\mathbf{y}) + H_y(\mathbf{y})\right] + \eta_y(t)$$

$$P_\infty(\mathbf{y}) = \frac{1}{Z_y} \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

Equilibrium:
*genotypic free energy*

$$F_y = -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

$$= -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y H_y(\mathbf{y})} \left[Z_x(\mathbf{y})\right]^{\frac{\beta_y}{\beta_x}}$$

$$\left\langle \nabla_y H(\mathbf{x},\mathbf{y}) \right\rangle_{\mathbf{x}|\mathbf{y}} = \frac{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} \nabla_y H(\mathbf{x},\mathbf{y})}{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}}$$

$$= -\frac{1}{\beta_x} \nabla_y \log \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} = \nabla_y \underbrace{F_x(\mathbf{y})}_{\textit{free energy}}$$

hence

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \left[F_x(\mathbf{y}) + H_y(\mathbf{y})\right] + \eta_y(t)$$

$$P_\infty(\mathbf{y}) = \frac{1}{Z_y} \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

Equilibrium:
*genotypic free energy*

$$F_y = -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

$$= -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y H_y(\mathbf{y})} \left[Z_x(\mathbf{y})\right]^{\frac{\beta_y}{\beta_x}}$$

$$\left\langle \nabla_y H(\mathbf{x},\mathbf{y}) \right\rangle_{\mathbf{x}|\mathbf{y}} = \frac{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} \nabla_y H(\mathbf{x},\mathbf{y})}{\int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}}$$

$$= -\frac{1}{\beta_x} \nabla_y \log \int \mathrm{d}\mathbf{x}\, \mathrm{e}^{-\beta_x \left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]} = \nabla_y \underbrace{F_x(\mathbf{y})}_{\textit{free energy}}$$

hence

$$\tau_y \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y} = -\nabla_y \left[F_x(\mathbf{y}) + H_y(\mathbf{y})\right] + \eta_y(t)$$

$$P_\infty(\mathbf{y}) = \frac{1}{Z_y} \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

Equilibrium:
*genotypic free energy*

$$F_y = -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y \left[F_x(\mathbf{y})+H_y(\mathbf{y})\right]}$$

$$= -\frac{1}{\beta_y} \log \int \mathrm{d}\mathbf{y}\, \mathrm{e}^{-\beta_y H_y(\mathbf{y})} \left[Z_x(\mathbf{y})\right]^{\frac{\beta_y}{\beta_x}}$$

# Replica method

$n = \beta_y/\beta_x$, $\beta_x = \beta$:

$$F_y = -\frac{1}{n\beta}\log\int \mathrm{d}\mathbf{y}\; \mathrm{e}^{-\beta n H_y(\mathbf{y})}\left[Z_x(\mathbf{y})\right]^n, \qquad Z_x(\mathbf{y}) = \int \mathrm{d}\mathbf{x}\; \mathrm{e}^{-\beta\left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}$$

(i) evaluate $Z^n$ for *integer* $n$
(ii) compute $F_y$
(iii) analytical continuation to non-integer $n$

$$F_y = -\frac{1}{n\beta}\log\int \mathrm{d}\mathbf{x}^1\ldots\mathrm{d}\mathbf{x}^n\; \mathrm{e}^{-\beta\sum_{\alpha=1}^n H_x(\mathbf{x}^\alpha)-\beta H_{\mathrm{eff}}(\mathbf{x}^1,\ldots,\mathbf{x}^n)}$$

$$H_{\mathrm{eff}}(\mathbf{x}^1,\ldots,\mathbf{x}^n) = -\frac{1}{\beta}\log\int \mathrm{d}\mathbf{y}\; \mathrm{e}^{-\beta n H_y(\mathbf{y})-\beta\sum_{\alpha=1}^n H(\mathbf{x}^\alpha,\mathbf{y})}$$

- $n\to 0$: random genotypes
- $n\to 1$: annealed averages (as if timescales same)
- $n>2$: first order transitions ...

# Replica method

$n = \beta_y/\beta_x$, $\beta_x = \beta$:

$$F_y = -\frac{1}{n\beta}\log\int \mathrm{d}\mathbf{y}\; \mathrm{e}^{-\beta n H_y(\mathbf{y})}\left[Z_x(\mathbf{y})\right]^n, \qquad Z_x(\mathbf{y}) = \int \mathrm{d}\mathbf{x}\; \mathrm{e}^{-\beta\left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}$$

(i) evaluate $Z^n$ for *integer* $n$
(ii) compute $F_y$
(iii) analytical continuation to non-integer $n$

$$F_y = -\frac{1}{n\beta}\log\int \mathrm{d}\mathbf{x}^1\ldots\mathrm{d}\mathbf{x}^n\; \mathrm{e}^{-\beta\sum_{\alpha=1}^n H_x(\mathbf{x}^\alpha)-\beta H_{\mathrm{eff}}(\mathbf{x}^1,\ldots,\mathbf{x}^n)}$$

$$H_{\mathrm{eff}}(\mathbf{x}^1,\ldots,\mathbf{x}^n) = -\frac{1}{\beta}\log\int \mathrm{d}\mathbf{y}\; \mathrm{e}^{-\beta n H_y(\mathbf{y})-\beta\sum_{\alpha=1}^n H(\mathbf{x}^\alpha,\mathbf{y})}$$

- $n\to 0$: random genotypes
- $n\to 1$: annealed averages (as if timescales same)
- $n>2$: first order transitions ...

# Replica method

$n = \beta_y/\beta_x$, $\beta_x = \beta$:

$$F_y = -\frac{1}{n\beta}\log\int \mathrm{d}\mathbf{y}\; \mathrm{e}^{-\beta n H_y(\mathbf{y})}\left[Z_x(\mathbf{y})\right]^n, \qquad Z_x(\mathbf{y}) = \int \mathrm{d}\mathbf{x}\; \mathrm{e}^{-\beta\left[H(\mathbf{x},\mathbf{y})+H_x(\mathbf{x})\right]}$$

(i) evaluate $Z^n$ for *integer* $n$
(ii) compute $F_y$
(iii) analytical continuation to non-integer $n$

$$F_y = -\frac{1}{n\beta}\log\int \mathrm{d}\mathbf{x}^1\ldots\mathrm{d}\mathbf{x}^n\; \mathrm{e}^{-\beta\sum_{\alpha=1}^n H_x(\mathbf{x}^\alpha)-\beta H_{\mathrm{eff}}(\mathbf{x}^1,\ldots,\mathbf{x}^n)}$$

$$H_{\mathrm{eff}}(\mathbf{x}^1,\ldots,\mathbf{x}^n) = -\frac{1}{\beta}\log\int \mathrm{d}\mathbf{y}\; \mathrm{e}^{-\beta n H_y(\mathbf{y})-\beta\sum_{\alpha=1}^n H(\mathbf{x}^\alpha,\mathbf{y})}$$

- $n\to 0$: random genotypes
- $n\to 1$: annealed averages (as if timescales same)
- $n>2$: first order transitions ...

1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-$n$ replica approach to biological evolution

4 Open questions

5 Summary

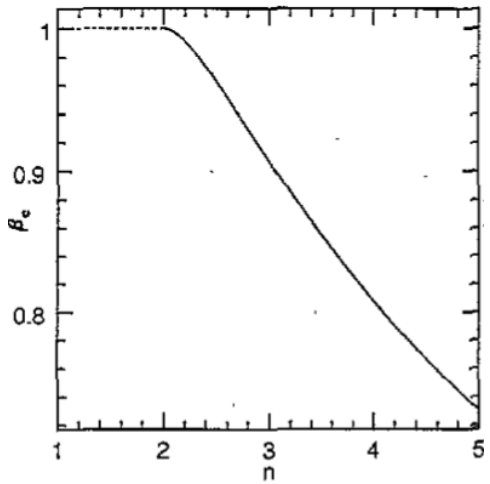
# Open questions

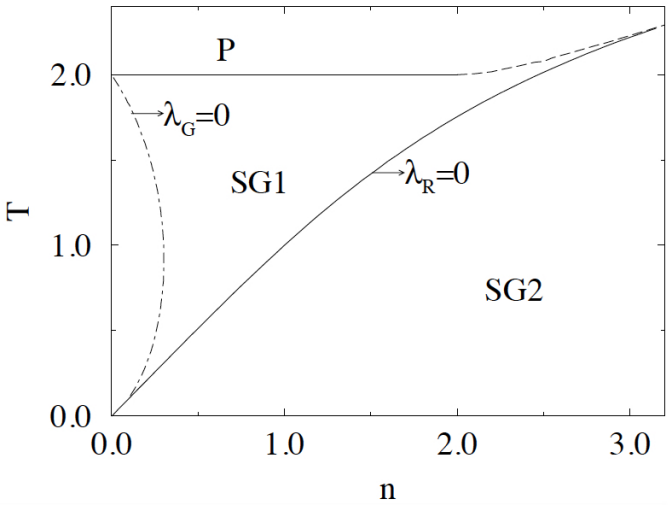
- discrete slow variables:

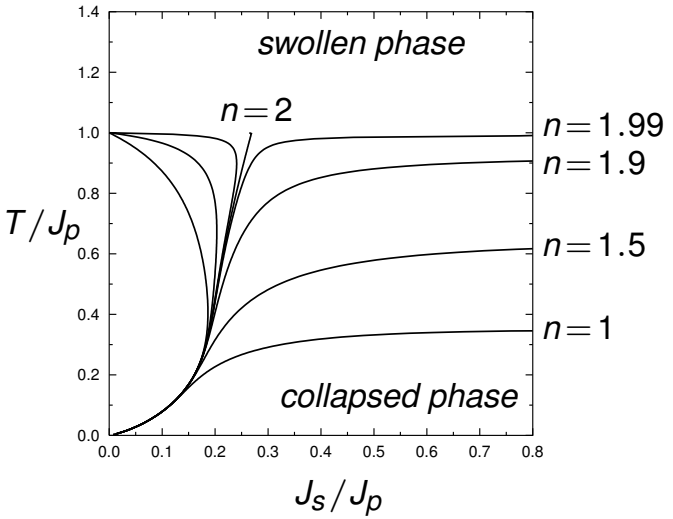
  nontrivial to define stochastic dynamics such that
  slow Hamiltonian is free energy of the fast degrees of freedom ...

- $n = 2$: boundary between 1st and 2nd order transitions
  in coupled dynamics models with adiabatically separated time scales

  what is special about $T_x / T_y = 2$?



1 Introduction
- Biological evolution as a modelling challenge
- Heterogeneity in physics versus biology

2 Inspiration from statistical physics
- Ising spin systems with Hebbian bond evolution
- XY spins and added layers of complexity
- Link with the Parisi scheme

3 Application to biology
- Self-programming in neural systems
- Genesis of protein structure
- Finite-$n$ replica approach to biological evolution

4 Open questions

5 Summary

# Summary

- Evolution: coupled dynamics of fast and slow degrees of freedom (phenotype and genotype)
- Assume both minimise stochastically the *same* energy function ('fitness' + constraints + utility)
- In the regime of adiabatically separated timescales: finite $n$ replica theories in equilibrium, with $n = T_{\mathrm{fast}}/T_{\mathrm{slow}}$
- *Models analytically solvable, expressions for heterogeneity distribution*
- For $n > 2$ (strong coupling between slow and fast variables): first order phase transitions