Introduction to the replica method

Yoshiyuki Kabashima
The Institute for Physics of Intelligence
& Department of Physics
The University of Tokyo

Back ground and motivation

- Many problems in information science have similarity to many-body problems in physics
- However, methods/styles of analysis developed in the two disciplines look quite different
- This implies that importing/exporting notions and techniques from/to the other may lead to novel findings in one field

Purpose

- Having such a perspective, we here introduce a physics-based technique developed for analyzing disordered many-body problems, which is now becoming popular more in information science.
  - ***Replica method***

Outline

- Part I: Structural similarity between physics of disordered systems and information science
  - Random energy model ← Physics
  - Error correcting codes ← Information theory
  - Random k-SAT problem ← Theoretical computer science

- Part II: Demonstration of the replica calculation
  - Replica analysis of random energy model

PART I: STRUCTURAL SIMILARITY BETWEEN PHYSICS OF DISORDERED SYSTEMS AND INFORMATION SCIENCE

Similarity between physics and information science

- From now, we introduce three problems whose origins and back grounds are unrelated with one another.

- However, their mathematical structure and technical difficulty are very similar.

- After the introduction, we *formally* show how the replica method can potentially resolve the difficulty in conjunction with its mathematical faults.

1) Random energy model (REM)

- A toy model introduced by Derrida (1980)
  - For each state $\boldsymbol{\tau} \in \{+1, -1\}^N$, assign an energy value $E(\boldsymbol{\tau})$ *randomly* by i.i.d. sampling from

$$P(E) = \frac{1}{\sqrt{N\pi}} \exp\left[-\frac{E^2}{N}\right]$$

  - Defines energy function modeling complicated interactions cf) spin glass, glasses, polymers, proteins, etc

$$H(\mathbf{s}|\mathbf{E}) = \sum_{\tau} E(\boldsymbol{\tau})\delta(\mathbf{s}, \boldsymbol{\tau})$$

- Problem: Evaluate macroscopic quantities for the canonical distribution for large system limit $N \to \infty$

$$P_\beta(\mathbf{s}|\mathbf{E}) = \frac{1}{Z(\beta|\mathbf{E})} \exp[-\beta H(\mathbf{s}|\mathbf{E})]$$

Macroscopic quantities

- Internal energy/free energy/entropy (densities)

$$u = \frac{1}{N}\sum_{\mathbf{s}} H\left(\mathbf{s}|\mathbf{E}\right)P_\beta\left(\mathbf{s}|\mathbf{E}\right)$$

(internal energy)

$$f = -\frac{1}{N\beta}\ln Z\left(\beta|\mathbf{E}\right)$$

(free energy)

$$s = -\frac{1}{N}\sum_{\mathbf{s}} P_\beta\left(\mathbf{s}|\mathbf{E}\right)\ln P_\beta\left(\mathbf{s}|\mathbf{E}\right)$$

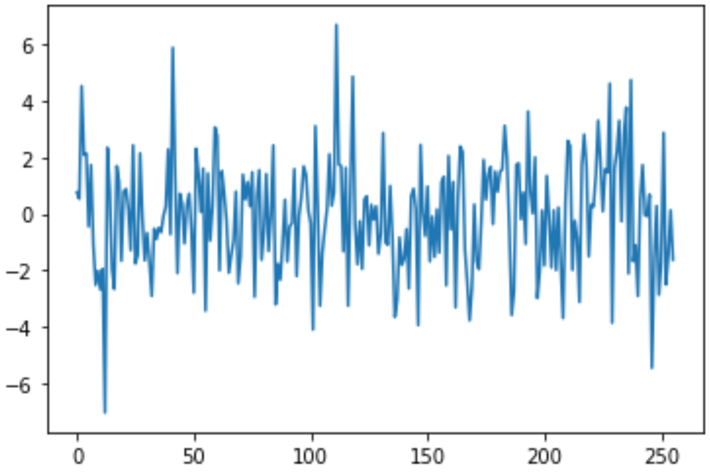
(entropy)

- Obviously, they depend on each sample of $\mathbf{E}$

$N = 8$

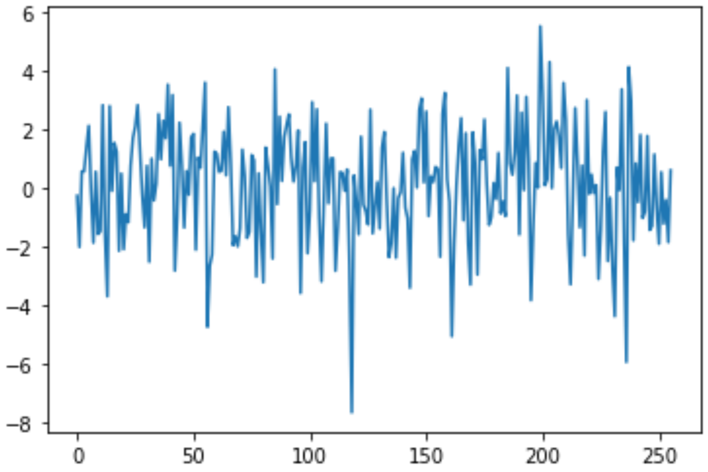
$\mathbf{E}(\boldsymbol{\tau})$

Sample 1



$\boldsymbol{\tau} \in \{+1,-1\}^8$

Sample 2



$\boldsymbol{\tau} \in \{+1,-1\}^8$

Self-averaging property

- However, for $N \to \infty$, the macroscopic quantities of typical samples of REM, converge to their expectations as

$$u \to [u]_{\mathbf{E}} \qquad f \to [f]_{\mathbf{E}} \qquad s \to [s]_{\mathbf{E}}$$

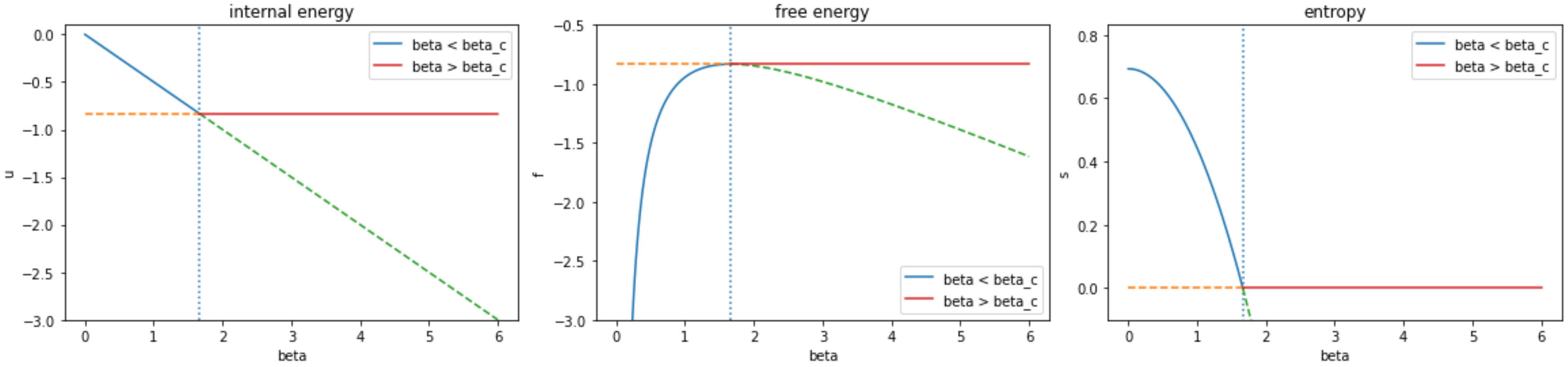
- Typical samples
  - For $\forall \epsilon > 0$, samples that satisfy

$$\left| \underbrace{2^{-N} \sum_{\tau} -\ln P\left(E(\tau)\right)}_{\text{Empirical information (per state) of } \mathbf{E}} - \underbrace{\left(\frac{1}{2} \ln\left(N\pi e\right) \right)}_{\text{Entropy (density) of } \mathbf{E}} \right| < \epsilon$$

  - For $N \to \infty$, the fraction of typical samples converges to unity

Phase transition

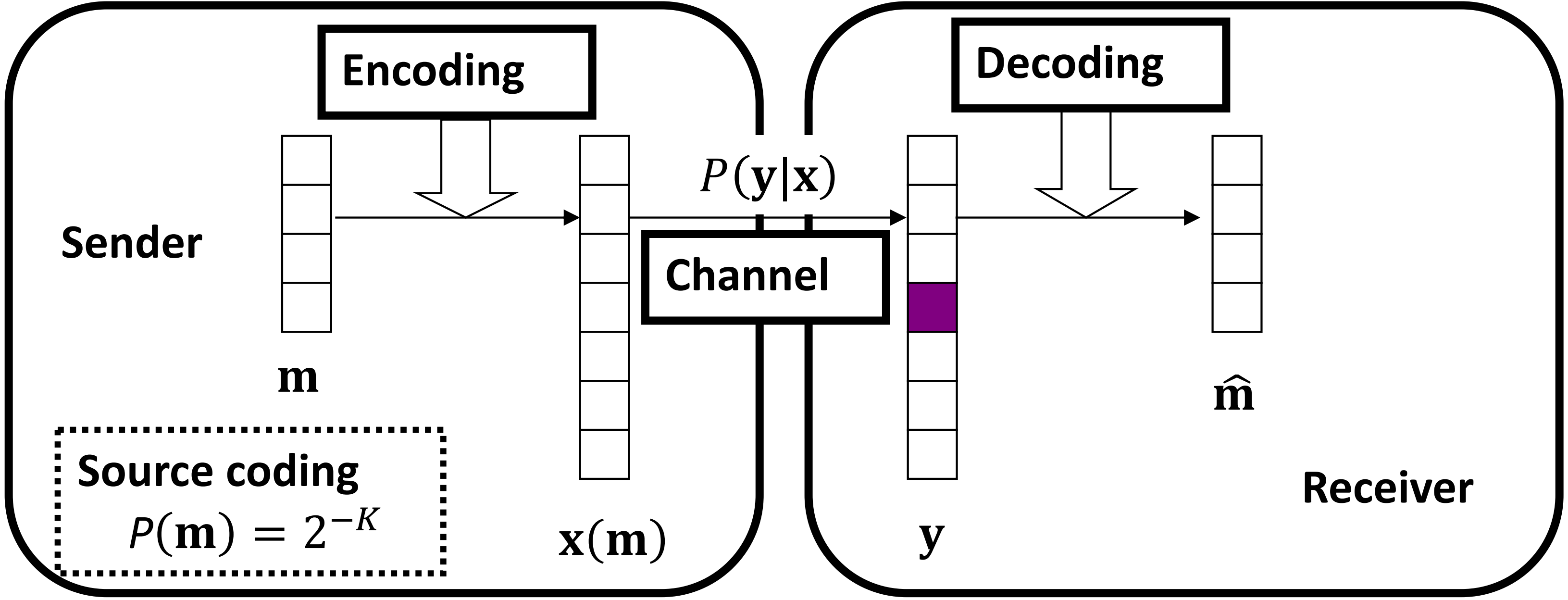
- As long as N is finite, $[u]_{\mathbf{E}}$, $[f]_{\mathbf{E}}$, $[s]_{\mathbf{E}}$ are analytic with respect to inverse temperature β
- However, for $N \to \infty$, the analyticity of these functions is broken at $\beta_c = 2\sqrt{\ln 2}$
  - ***Phase (freezing) transition*** (details are shown in 2nd part)
  - Explains the generality of *"frozen behavior"* in low temperatures in complex systems



2) Error correcting codes

- Shannon (1948)
  - Reliable communication via noisy channel
  - Channel coding (error correcting code): Original message $\mathbf{m} \in \{+1,-1\}^K$ is encoded into a redundant expression (codeword) $\mathbf{x}(\mathbf{m}) \in \{+1,-1\}^N$

Communication under code rate $R = \frac{K}{N}$



Decoding problem

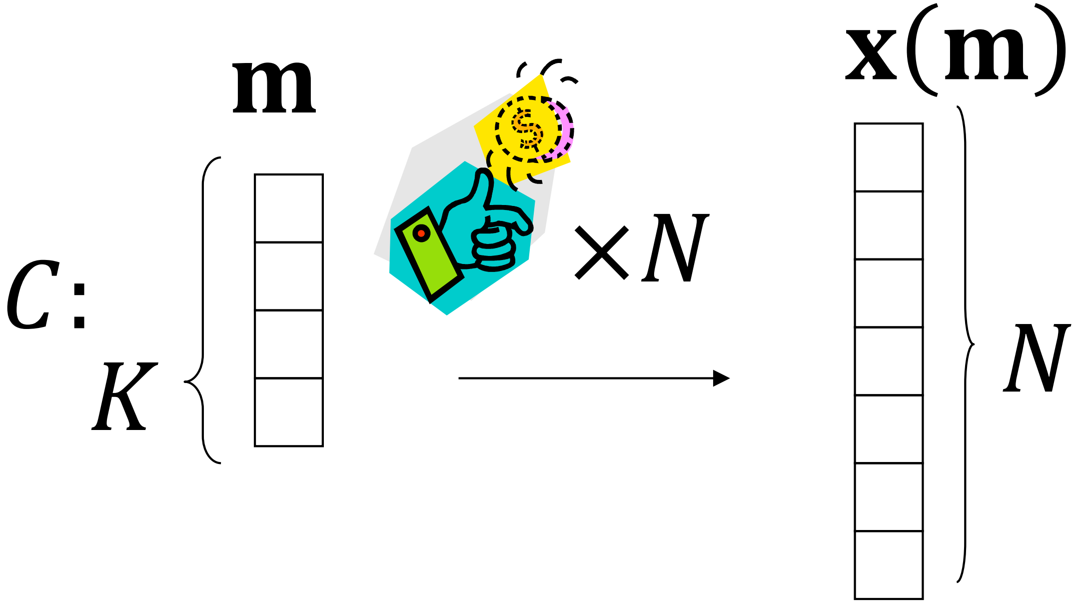
- Decoding:
  Infer the original message $\mathbf{m}$ from a received (noisy) codeword $\mathbf{y}$
  - Bayes' theorem

$$P(m|y) = \frac{P\big(y|x(m)\big)P(m)}{\sum_{m'} P\big(y|x(m')\big)P(m')} = \frac{P\big(y|x(m)\big)}{\sum_{m'} P\big(y|x(m')\big)}$$

- Problem: Under what condition, is the original message correctly decodable for $N \to \infty$?

Random code ensemble (RCE)

- Construct the coding $C\colon \mathbf{m} \to \mathbf{x}(\mathbf{m})$ by fair coin-tossing

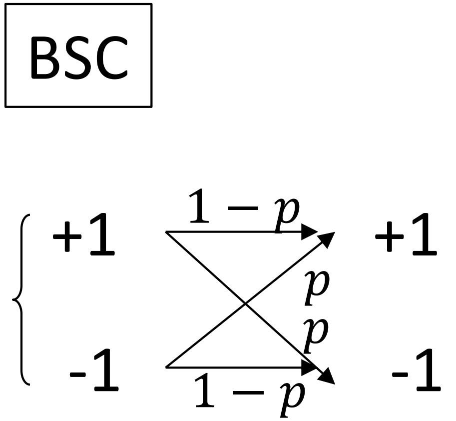


  - Requires $O(N\times 2^{K})$ storage space for keeping a code book for representing C
  - So, not practically in use

Channel coding theorem (for BSC; for simplicity)

- However, Shannon (1948) showed that RCE exhibits the best possible error correction ability
  - Useful baseline for assessing performance of practical codes

- For instance, for binary symmetric channel (BSC), the probability of decoding failure of typical samples of RCE becomes arbitrary small as $N \to \infty$, if code rate satisfies



$$R(= K / N) < 1 + p\log_2 p + (1-p)\log_2(1-p)$$

and no other codes achieves this performance

Similarity to REM

- Depends on *pre-determined (quenched)* randomness
  - Energy function (REM), codebook and noise (RCE)

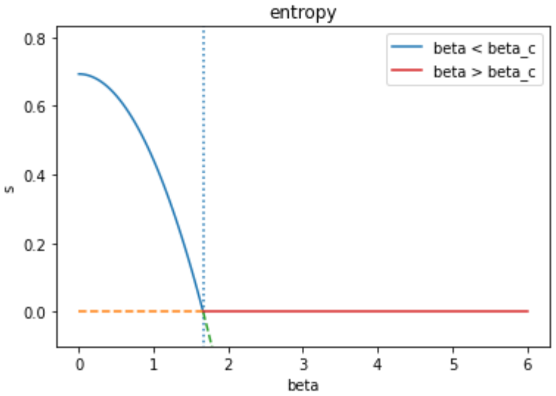
REM

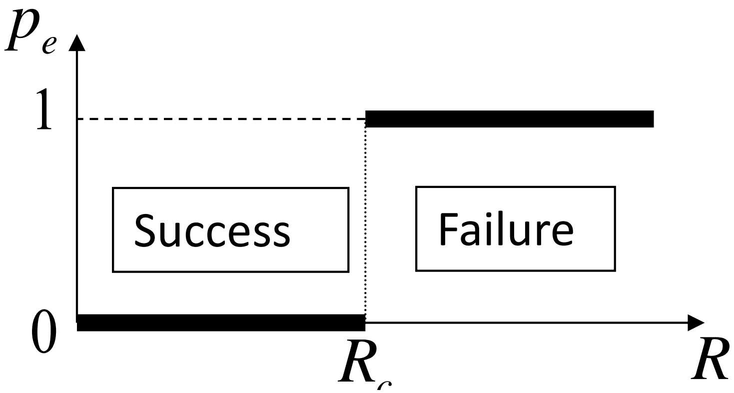
$$P_\beta(\mathbf{s}|\mathbf{E}) = \frac{1}{Z(\beta|\mathbf{E})}\exp\left(-\beta H(\mathbf{s}|\mathbf{E})\right)$$

RCE

$$P(\mathbf{m}|\mathbf{y},C) = \frac{P(\mathbf{y}|\mathbf{x}(\mathbf{m},C))P(\mathbf{m})}{\sum_{\mathbf{m}'}P(\mathbf{y}|\mathbf{x}(\mathbf{m}',C))P(\mathbf{m}')}$$

- Macroscopic quantities *typically converge to deterministic values in the limit of* $N \to \infty$
- Breaking of analyticity *(phase transition)*





3) Random k-SAT problems

- k-SAT problem
  - Determine if the variables of a given k-CNF formula has at least one assignment of Boolean variables that makes the formula evaluate to TRUE(=1)

Boolean variables $\mathbf{x} \in \{0,1\}^N$

k-clause A tuple of at most k Boolean variables or their negations connected by "or"

$$C_1(\mathbf{x}) = x_2 \vee x_5 \vee \overline{x_7} \qquad C_2(\mathbf{x}) = x_1 \vee x_4 \vee x_9$$

$$C_3(\mathbf{x}) = \overline{x_3} \vee \overline{x_4} \vee x_7 \quad \cdots$$

k-conjunctive normal form (k-CNF) A tuple of k-clauses connected by "and"

$$F(\mathbf{x}|\mathbf{C}) = C_1(\mathbf{x}) \wedge C_2(\mathbf{x}) \wedge \cdots \wedge C_M(\mathbf{x})$$

3) Random k-SAT problems

- k-SAT problem
  - Determine if the variables of a given k-CNF formula has at least one assignment of Boolean variables that makes the formula evaluate to TRUE(=1)

Boolean variables

$$\mathbf{x} \in \{0,1\}^N$$

Has an important status in computational complexity theory (standard form of NP-complete class)

$$F\left(\mathbf{x}\middle|\mathbf{C}\right) = C_1\left(\mathbf{x}\right) \wedge C_2\left(\mathbf{x}\right) \wedge \cdots \wedge C_M\left(\mathbf{x}\right)$$

SAT/UNSAT transition

- Suppose a situation

Clauses: #

$$N, M \gg 1,\ \alpha = \frac{M}{N} \sim O(1)$$

Boolean variables: #

- The fraction of k-CNF formulas that have SAT solutions drastically changes at a critical ratio $\alpha_c(k)$
  - $\alpha_c(1) = 0, \alpha_c(2) = 1$ ← Theoretical
  - $\alpha_c(3) = 4.2\cdots$ ← Experimental

Monasson et al,
Nature 400, 133 (1999)

2+p-SAT

$$(N = 10^3 \sim 10^4)$$

$$k = 3 \rightarrow \alpha_c(k = 3) = 4.2\cdots$$

Stat. mech. expression of k-SAT

Binary-bipolar transformation

$$x_i \in \{0,1\} \rightarrow s_i = (-1)^{x_i} = 1 - 2x_i \in \{+1,-1\}$$

$$C(\mathbf{x}) = x_2 \vee x_5 \vee \overline{x_7} = 1 - \left(\frac{1-s_2}{2}\right)\left(\frac{1-s_5}{2}\right)\left(\frac{1+s_7}{2}\right)$$

Energy function = # Unsatisfied clauses

$$F(\mathbf{x}|\mathbf{C}) = C_1(\mathbf{x}) \wedge C_2(\mathbf{x}) \wedge \cdots \wedge C_M(\mathbf{x})$$

$$H(\mathbf{s}|\mathbf{c}) = \sum_{\mu=1}^{M} \prod_{i=1}^{k} \left(\frac{1 + c_{\mu\ell_i} s_{\ell_i}}{2}\right)$$

$$c_{\mu\ell_i} \in \{+1,-1\}$$

Affirmation/negation in μ-th clause

Stat. mech. expression of k-SAT

- SAT ⟺ min. energy = 0

Unsat. clauses

$$F\left(\exists \mathbf{x} \middle| \mathbf{C}\right) = 1 \quad \Longleftrightarrow \quad \min_{\mathbf{s}} \left\{ H\left(\mathbf{s} \middle| \mathbf{c}\right) \right\} = 0$$

- Min. energy = free energy for $\beta \to \infty$

$$\min_{\mathbf{s}} \left\{ H\left(\mathbf{s} \middle| \mathbf{c}\right) \right\} = \lim_{\beta \to \infty} -\frac{1}{\beta} \ln \left(\sum_{\mathbf{s}} \exp\left(-\beta H\left(\mathbf{s} \middle| \mathbf{c}\right) \right) \right)$$

$$= \lim_{\beta \to \infty} -\frac{1}{\beta} \ln Z\left(\beta \middle| \mathbf{c}\right)$$

Similarity to REM

- Depends on *pre-determined (quenched)* randomness
  - Energy function (REM), random k-CNF (k-SAT)

REM

$$P_\beta(\mathbf{s}|\mathbf{E}) = \frac{1}{Z(\beta|\mathbf{E})}\exp\left(-\beta H(\mathbf{s}|\mathbf{E})\right)$$

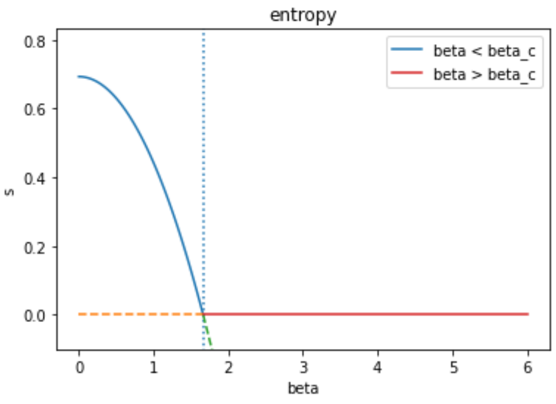
k-SAT

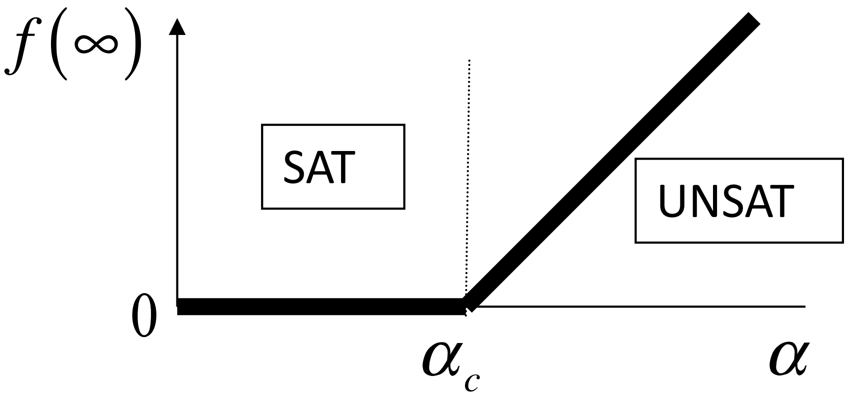
$$P_\beta(\mathbf{s}|\mathbf{c}) = \frac{\exp\left(-\beta H(\mathbf{s}|\mathbf{c})\right)}{Z(\beta|\mathbf{c})}$$

- Macroscopic quantities *typically converge to deterministic values in the limit of $N \to \infty$*

$$f(\beta|\mathbf{c}) = -\frac{1}{N\beta}\ln Z(\beta|\mathbf{c}) \to \left[f(\beta|\mathbf{c})\right]_\mathbf{c} = -\frac{1}{N\beta}\left[\ln Z(\beta|\mathbf{c})\right]_\mathbf{c}$$

- Breaking of analyticity *(phase transition)*





Unified perspective

- Common structure of the three examples
  - Conditional distribution

$$P_\beta(\mathbf{s}|\mathbf{E})=\frac{1}{Z(\beta|\mathbf{E})}\exp\left(-\beta H(\mathbf{s}|\mathbf{E})\right)$$

- The key to solve the problems is the assessment of the average free energy

$$f(\beta)=-\lim_{N\to\infty}\frac{1}{N\beta}\left[\ln Z(\beta|\mathbf{E})\right]_{\mathbf{E}}=-\lim_{N\to\infty}\frac{1}{N\beta}\sum_{\mathbf{E}}P(\mathbf{E})\ln\left(\sum_{\mathbf{s}}\exp\left(-\beta H(\mathbf{s}|\mathbf{E})\right)\right)$$

  - Once the free energy is obtained, other quantities can be assessed from it

$$u(\beta)=\frac{\partial}{\partial\beta}\left(\beta f(\beta)\right) \qquad s(\beta)=\beta\left(u(\beta)-f(\beta)\right)$$

Technical difficulty

- Unfortunately, averaging ``$\ln Z$'' is difficult to perform in general

$$\left[\ln Z\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}} = \sum_{\mathbf{E}} P\left(\mathbf{E}\right)\ln \underbrace{\left(\sum_{\mathbf{s}}\exp\left(-\beta H\left(\mathbf{s}|\mathbf{E}\right)\right)\right)}_{\text{Generally produces complicated dependence among components of } \mathbf{E}}$$

Moment function

- On the other hand, averaging ``Z^n'' is relatively easy to perform for natural numbers $n = 1,2,\ldots \in \mathbb{N}$ using the expansion formula as

$$\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} = \sum_{\mathbf{E}} P(\mathbf{E}) \left(\sum_{\mathbf{s}} \exp\left(-\beta H(\mathbf{s}|\mathbf{E})\right) \right)^n$$

$$= \sum_{\mathbf{E}} P(\mathbf{E}) \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n} \exp\left(-\beta \sum_{a=1}^{n} H\left(\mathbf{s}^a \middle| \mathbf{E}\right) \right)$$

$$= \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n} \boxed{\sum_{\mathbf{E}} P(\mathbf{E}) \exp\left(-\beta \sum_{a=1}^{n} H\left(\mathbf{s}^a \middle| \mathbf{E}\right) \right)}$$

$$= \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n} \exp\left(-\beta \mathcal{H}_n\left(\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n;\beta\right)\right)$$

$$\triangleq \exp\left(-\beta \mathcal{H}_n\left(\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n;\beta\right)\right)$$

Effective Boltzmann weight

Replica method

1. Evaluate $[Z^n(\beta|\mathbf{E})]_{\mathbf{E}}$ for $n = 1,2,\ldots \in \mathbb{N}$ as a function of n
2. Analytically continue the obtained functional expression from $n = 1,2,\ldots \in \mathbb{N}$ to real numbers $n \in \mathbb{R}$

$$\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} = \exp\left(N\phi(n,\beta)\right)$$

$$\left(n \in \mathbb{N} \to n \in \mathbb{R}\right)$$

3. Evaluate the average free energy using an identity (replica trick) as

$$\frac{1}{N}\left[\ln Z(\beta|\mathbf{E})\right]_{\mathbf{E}} = \frac{1}{N}\lim_{n\to 0}\frac{\partial}{\partial n}\ln\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}}$$

$$= \lim_{n\to 0}\frac{\partial}{\partial n}\phi(n,\beta)$$

Remark (I)

- The idea of the “replica trick” has a long history, dating back at least to 1930s, although it had not been so popular until its application to the spin glass problem in 1970s.

Remark (II)

- After the employment of the expansion formula, n dynamical variables come out. They are regarded as representing n copies (replicas) of the original system that share the same predetermined randomness. This is the origin of the name of the "replica method".

$$Z^n\left(\beta \middle| \mathbf{E}\right) = \left(\sum_{\mathbf{s}} \exp\left(-\beta H\left(\mathbf{s} \middle| \mathbf{E}\right)\right) \right)^n$$

$$= \sum_{\underbrace{\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^n}_{n \text{ replicas}}} \exp\left(-\beta \sum_{a=1}^{n} H\left(\mathbf{s}^a \middle| \mathbf{E}\right) \right)$$

Predetermined randomness

Remark (III)

- After performing the "configurational (quenched)" average with respect to the predetermined randomness, the problem is reduced to the computation of the partition function of an effective pure system of the n replicas.
- Therefore, one can exploit standard statistical mechanics techniques, which were developed for pure systems, for the reduced problem

$$\left[Z^n\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}} = \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n}\sum_{\mathbf{E}} P\left(\mathbf{E}\right)\exp\left(-\beta\sum_{a=1}^{n} H\left(\mathbf{s}^a|\mathbf{E}\right)\right)$$

$$= \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n}\exp\left(-\beta\underbrace{\mathcal{H}_n\left(\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n;\beta\right)}\right)$$

- Effective Hamiltonian for n replica system
- No randomness

Mathematical faults of the replica method

- As shown, there are many problems to which the replica method can be *potentially* applied. Actually, it has yielded a number of nontrivial findings in various fields.
  - Spin glasses, polymers, neural networks, machine learning, error correcting codes, SAT problems, wireless communication, signal processing, etc
- However, there are two intrinsic open problems in the replica method, which makes its status “non-rigorous heuristics”.
- Before proceeding to technical details, we wrap up the first part mentioning the two problems.
  - Nevertheless, we have to say that there is no known example to which the replica method leads to wrong results by appropriately taking into account the “replica symmetry breaking” if necessary.

Two intrinsic problems of the replica method (I)

- Analytical continuation from natural numbers $n = 1,2, \ldots \in \mathbb{N}$ to real (or complex) numbers $n \in \mathbb{R}$ (or $\mathbb{C}$) cannot be defined uniquely in general.
  - Simple example

$$\phi(n,\beta), \quad \tilde{\phi}_a(n,\beta) = \phi(n,\beta) + a\sin(\pi n)$$

$$\begin{cases} \phi(n,\beta) = \tilde{\phi}_a(n,\beta) \text{ for } \forall a \text{ and } n = 1,2,\ldots \\ \lim_{n\to 0} \frac{\partial}{\partial n}\phi(n,\beta) \neq \lim_{n\to 0} \frac{\partial}{\partial n}\tilde{\phi}_a(n,\beta) \text{ if } a \neq 0 \end{cases}$$

  - This possibility is excluded if $\left(\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}}\right)^{1/N} \le \exp(Cn)$holds (Carlson's theorem). However, this does not hold in many problems.

Two intrinsic problems of the replica method (II)

- In practice, we need to swap the two limit operations for evaluating the effective partition function in most cases

$$\lim_{N\to\infty}\frac{1}{N}\left[\ln Z\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}}=\lim_{N\to\infty}\frac{1}{N}\lim_{n\to 0}\frac{\partial}{\partial n}\ln\left(\left[Z^n\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}}\right)\qquad\text{(what we have to do)}$$

$$\to\lim_{n\to 0}\frac{\partial}{\partial n}\left(\lim_{N\to\infty}\frac{1}{N}\ln\left(\left[Z^n\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}}\right)\right)\qquad\text{(what we can do)}$$

- This can lead to a wrong result when breaking of analyticity with respect to n occurs in the limit of $N \to \infty$
- One can mathematically show that the analyticity breaking w.r.t. n actually occurs in certain systems
  - Nevertheless, the correct solution can still be found by taking into account the “replica symmetry breaking” (Ogure and YK, PTP 111, 661 (2004); JSTAT (2009) P03010, P05011)

Summary of part I

- Various problems from physics and information science can be formulated in the form of conditional distributions (or Bayes theorem)

- Assessment of the configurational average of the logarithm of the partition function with respect to the predetermined randomness is the key to analyzing the typical property of the objective systems.

- The replica method is a systematic technique to performing the configurational average, but the method itself is not mathematically justified yet.

PART II: DEMONSTRATION OF THE REPLICA CALCULATION

Purpose

- Illustration of the replica method by applying it to a simple problem
  - random energy model (REM)

Outline

- Analysis of random energy model (REM) without using the replica method
- Replica analysis of REM

What is the replica method?

- A technique to evaluate general moment function $[Z^n(J)]_J$ $(n \in \mathbb{R})$ for disordered systems
- In many cases, used for evaluating $[\ln Z(J)]_J$
  - Replica trick

$$\left[\ln Z(\mathbf{J})\right]_{\mathbf{J}} = \lim_{n \to 0} \frac{\partial}{\partial n} \ln\left[Z^n(\mathbf{J})\right]_{\mathbf{J}} = \lim_{n \to 0} \frac{\left[Z^n(\mathbf{J})\right]_{\mathbf{J}} - 1}{n}$$

  - One can find its origin in mathematics
    - G.H. Hardy, Messenger Math. 58 (1929), 115.
    - G.H. Hardy, J.E. Littlewood and G. Polya, *Inequalities* (Campridge UP, 1934)

Sketch of RM

- Formula only valid for n=1,2,…

expansion

$$\left[Z^n(\mathbf{J})\right]_{\mathbf{J}} = \sum_{\mathbf{J}} P(\mathbf{J}) \left(\sum_{\mathbf{s}} e^{-\beta H(\mathbf{s}|\mathbf{J})} \right)^n = \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n} \sum_{\mathbf{J}} P(\mathbf{J}) e^{-\beta \sum_{a=1}^{n} H(\mathbf{s}^a|\mathbf{J})}$$

$$= \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n} e^{-\beta H_{\mathrm{eff}}\left(\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n;\beta\right)}$$

←Evaluate as a function of n

Analytical Continuation

$$\left[Z^n(\mathbf{J})\right]_{\mathbf{J}} \ \underline{\left(n \in \mathbb{N}\right)} \Rightarrow \left[Z^n(\mathbf{J})\right]_{\mathbf{J}} \ \underline{\left(n \in \mathbb{R}\right)}$$

Easy to evaluate

Hard to evaluate

Demonstration of RM

- Here, we demonstrate the actual computation of RM for the simplest spin glass model, *random energy model* (REM)

Random energy model (REM)

- A toy model introduced by Derrida (1980)
  - For each state $\boldsymbol{\tau} \in \{+1, -1\}^N$, assign an energy value $E(\boldsymbol{\tau})$ by i.i.d. sampling from

$$P(E) = \frac{1}{\sqrt{N\pi}} \exp\left[-\frac{E^2}{N}\right]$$

  - Energy function: modeling complicated interactions cf) spin glass, glasses, polymers, proteins, etc

$$H(\mathbf{s}|\mathbf{E}) = \sum_{\tau} E(\boldsymbol{\tau})\delta(\mathbf{s}, \boldsymbol{\tau})$$

- Problem: Evaluate macroscopic quantities for the canonical distribution for $N \to \infty$

$$P_\beta(\mathbf{s}|\mathbf{E}) = \frac{1}{Z(\beta|\mathbf{E})} \exp[-\beta H(\mathbf{s}|\mathbf{E})]$$

Macroscopic quantities

- Internal energy/free energy/entropy (densities)

$$u = \frac{1}{N}\sum_{\mathbf{s}} H(\mathbf{s}|\mathbf{E}) P_\beta(\mathbf{s}|\mathbf{E})$$

(internal energy)

$$f = -\frac{1}{N\beta}\ln Z(\beta|\mathbf{E})$$

(free energy)

$$s = -\frac{1}{N}\sum_{\mathbf{s}} P_\beta(\mathbf{s}|\mathbf{E}) \ln P_\beta(\mathbf{s}|\mathbf{E})$$

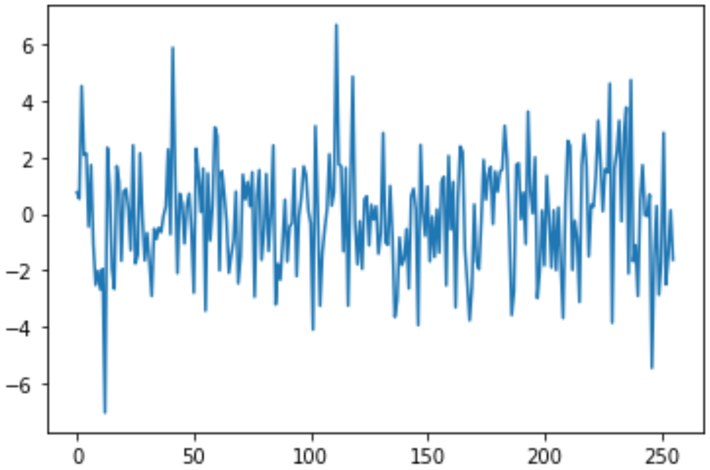
(entropy)

- Obviously, they depend on each sample of $\mathbf{E}$

$N = 8$

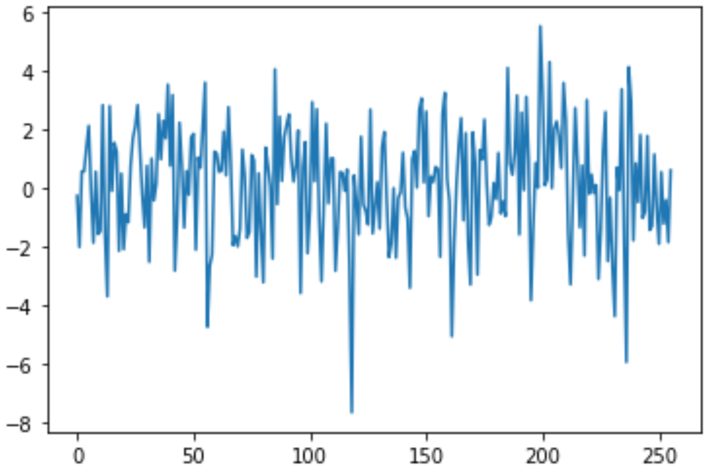
$\mathbf{E}(\boldsymbol{\tau})$

Sample 1



$\boldsymbol{\tau} \in \{+1, -1\}^8$

Sample 2



$\boldsymbol{\tau} \in \{+1, -1\}^8$

Analysis *without using* RM

- The number of states

$$\mathcal{N}\left(e\middle|\mathbf{E}\right)=\#\text{states whose energy } E\in\left[Ne,N\left(e+\delta e\right)\right]$$

- Its average and variance

$$\left[\mathcal{N}\left(e\middle|\mathbf{E}\right)\right]_{\mathbf{E}}=2^{N}\times P\left(E=Ne\right)\left(N\delta e\right)\sim\exp\left(N\left(\ln 2-e^{2}\right)\right)$$

$$\mathrm{var}\left[\mathcal{N}\left(e\middle|\mathbf{E}\right)\right]=2^{N}P\left(Ne\right)\left(N\delta e\right)\left(1-P\left(Ne\right)\left(N\delta e\right)\right)$$

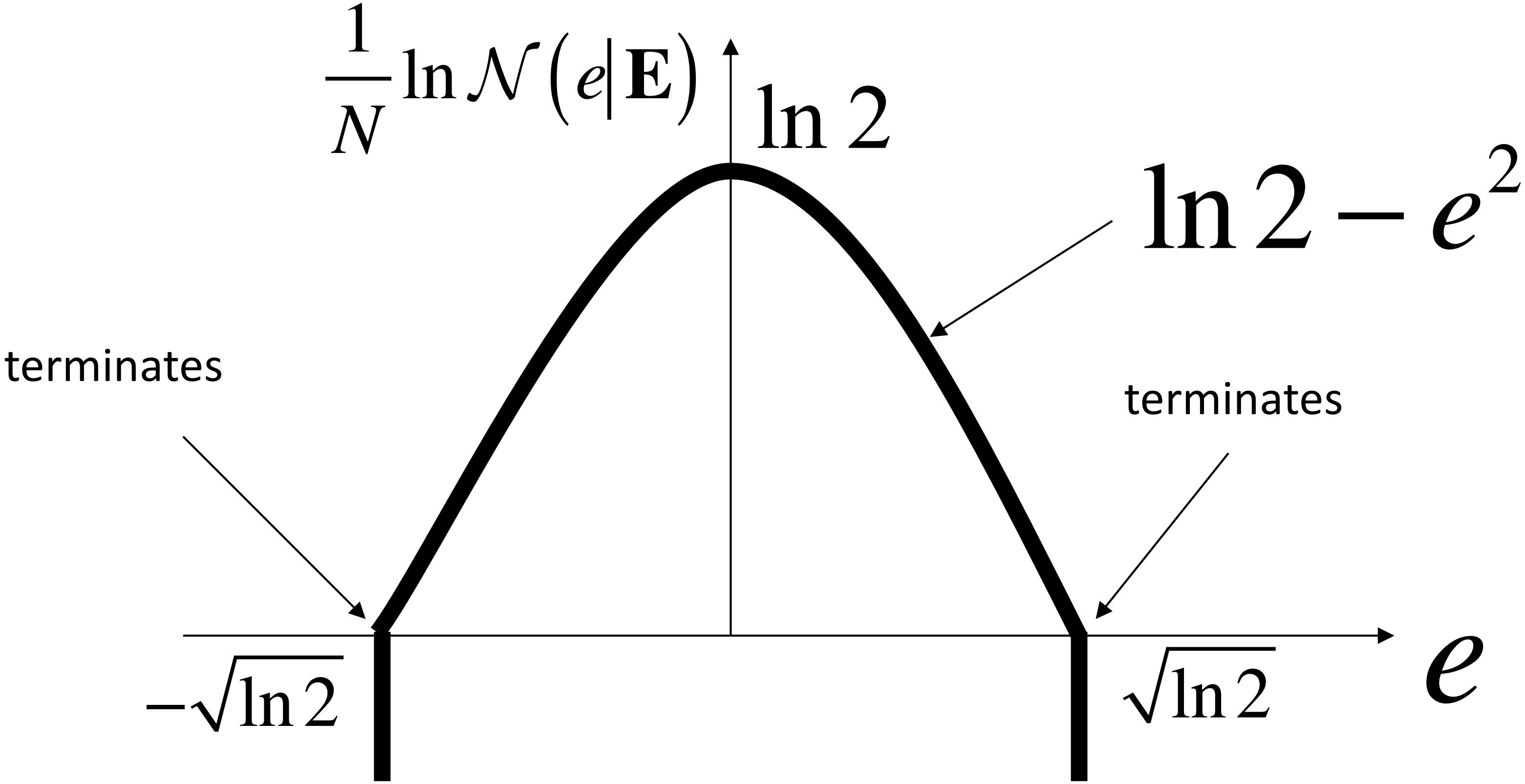
$$\sim\exp\left(N\left(\ln 2-e^{2}\right)\right)$$

  - $|e|<\sqrt{\ln 2}$: $\sqrt{\mathrm{var}\left[\mathcal{N}\left(e\middle|\mathbf{E}\right)\right]}\Big/\left[\mathcal{N}\left(e\middle|\mathbf{E}\right)\right]_{\mathbf{E}}\to 0$
  - $|e|>\sqrt{\ln 2}$: $\left[\mathcal{N}\left(e\middle|\mathbf{E}\right)\right]_{\mathbf{E}}\to 0$

⇨ For typical samples, no need for caring about statistical fluctuations

Schematic profile of #states

- For almost all realizations
  - Typical case analysis (Prob. → 1)
  - Not adequate for atypical (rare) cases

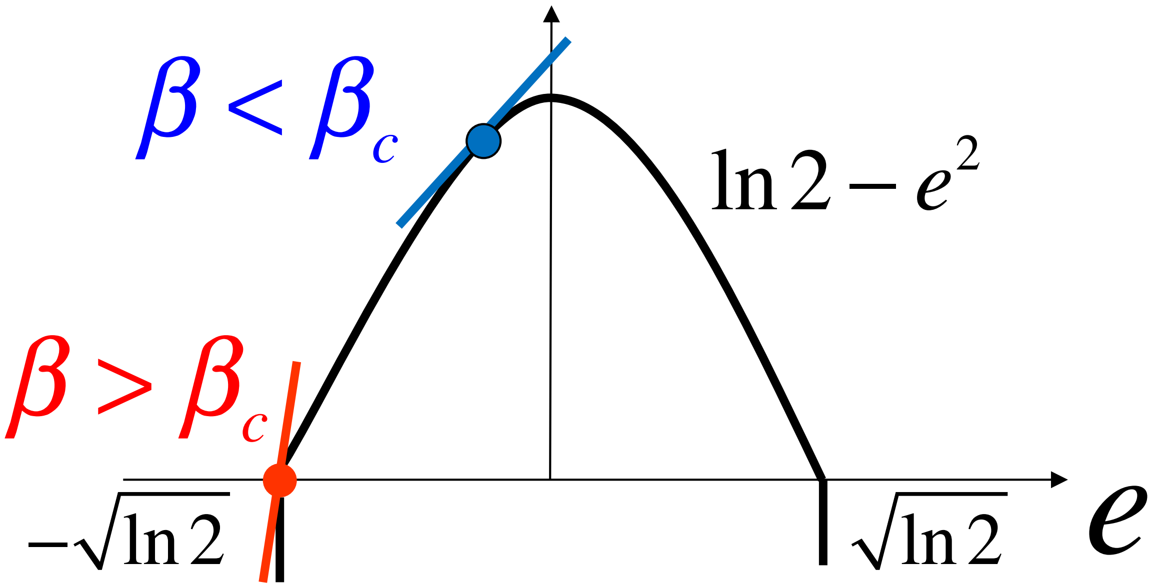


Evaluation by saddle point method

- Assessement by a single dominant contribution

$$\frac{1}{N}\ln Z\left(\beta\middle|\mathbf{E}\right) \simeq \frac{1}{N}\ln\left[\int d\left(Ne\right)e^{-\beta Ne}\left[\mathcal{N}\left(e\middle|\mathbf{E}\right)\right]_{\mathbf{E}}\right] \simeq \max_{e}\left\{-\beta e+\frac{1}{N}\ln\mathcal{N}\left(e\middle|\mathbf{E}\right)\right\}$$

$$= \max_{e\in\left[-\sqrt{\ln 2},\sqrt{\ln 2}\right]}\left\{-\beta e+\ln 2-e^2\right\} = \begin{cases} \frac{\beta^2}{4}+\ln 2, & \beta<\beta_c=2\sqrt{\ln 2} \\ \beta\sqrt{\ln 2}, & \beta>\beta_c=2\sqrt{\ln 2} \end{cases}$$



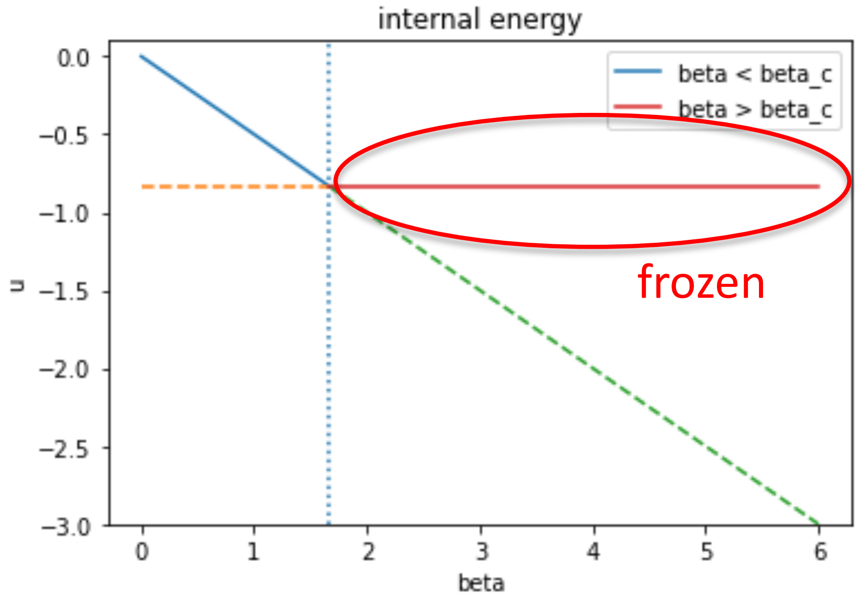
Correct result of REM

$$f(\beta) = -\frac{1}{N\beta}\ln Z(\beta|\mathbf{E}) = \begin{cases} -\frac{\beta}{4} - \frac{\ln 2}{\beta}, & \beta < \beta_c \left(= 2\sqrt{\ln 2}\right) \\ -\sqrt{\ln 2}, & \beta > \beta_c \end{cases}$$

$$u(\beta) = \frac{\partial\left(\beta f(\beta)\right)}{\partial\beta} = \begin{cases} -\frac{\beta}{2}, & \beta < \beta_c \\ -\sqrt{\ln 2}, & \beta > \beta_c \end{cases}$$

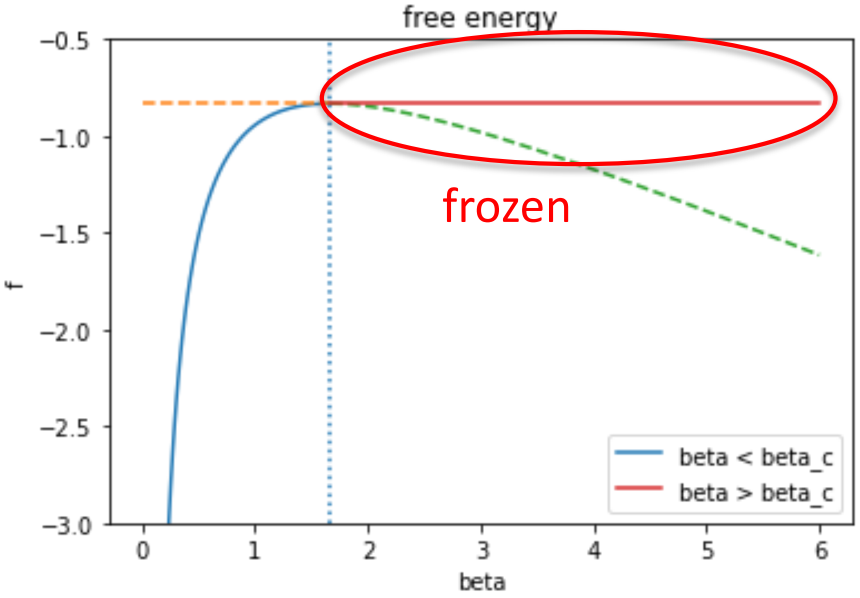
$$s(\beta) = \beta\left(u(\beta) - f(\beta)\right) = \begin{cases} -\frac{\beta^2}{4} + \ln 2, & \beta < \beta_c \\ 0, & \beta > \beta_c \end{cases}$$

Phase transition

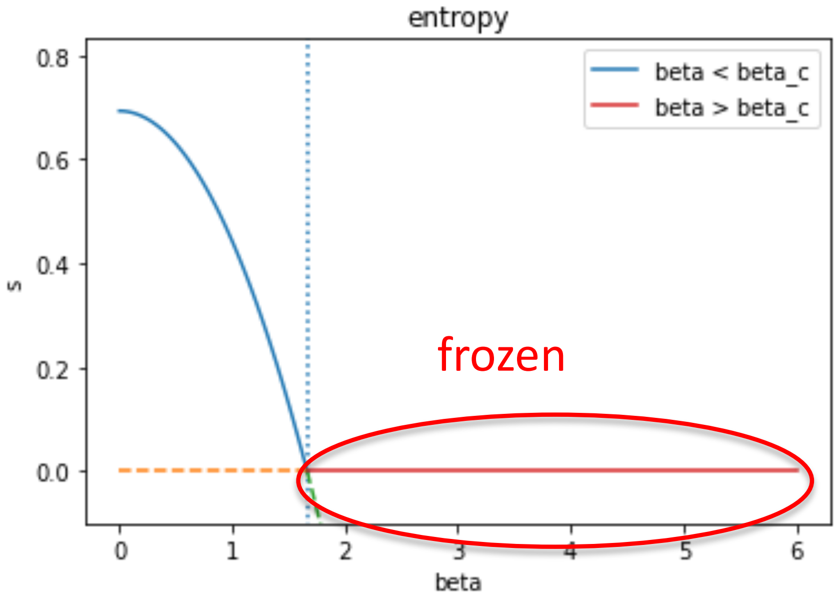


frozen

“Frozen” for $\beta > \beta_c$



frozen



frozen

Replica analysis of REM

$$f(\beta) = -\frac{1}{N\beta}\left[\ln Z(\beta|\mathbf{E})\right]_{\mathbf{E}}$$

$$= -\lim_{n\to 0}\frac{\partial}{\partial n}\lim_{N\to\infty}\frac{1}{N\beta}\ln\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}}$$

Analytical Continuation

$$\left[Z^n(\mathbf{J})\right]_{\mathbf{J}} \ \underline{(n \in \mathbb{N})} \quad \Rightarrow \quad \left[Z^n(\mathbf{J})\right]_{\mathbf{J}} \ \underline{(n \in \mathbb{R})}$$

Easy to evaluate

Hard to evaluate

Can RM reproduce the correct result?

Replication of partition function

- Partition function

$$Z\left(\beta\middle|\mathbf{E}\right)=\sum_{\mathbf{s}}\exp\left[-\beta E(\mathbf{s})\right]=\sum_{\mathbf{s}}\exp\left[-\beta\sum_{\tau}E(\tau)\delta(\mathbf{s},\tau)\right]$$

- Replication for $n = 1,2,\ldots \in \mathbb{N}$

$$Z^{n}\left(\beta\middle|\mathbf{E}\right)=\sum_{\mathbf{s}^{1},\mathbf{s}^{2},\ldots,\mathbf{s}^{n}}\exp\left[-\beta\sum_{\tau}E(\tau)\sum_{a=1}^{n}\delta\left(\mathbf{s}^{a},\tau\right)\right]$$

Key formula for configurational average

Average with respect to the energy of a single state τ

$$\left[\exp\left[-\beta E(\tau)\sum_{a=1}^{n}\delta\left(\mathbf{s}^{a},\tau\right)\right]\right]_{E(\tau)}$$

$$=\int\exp\left[-\beta E(\tau)\sum_{a=1}^{n}\delta\left(\mathbf{s}^{a},\tau\right)\right]\underbrace{\frac{e^{-\frac{(E(\tau))^{2}}{N}}}{\sqrt{N\pi}}}_{=P\left(E(\tau)\right)}dE(\tau)$$

$$=\exp\left[\frac{N}{4}\left(\beta\sum_{a=1}^{n}\delta\left(\mathbf{s}^{a},\tau\right)\right)^{2}\right]$$

Average of replicated Boltzmann factor

- For a fixed set of $\mathbf{s}^1, \mathbf{s}^2, \dots, \mathbf{s}^n$, the average of the replicated Boltzmann factor is labeled by a partition of n

$$\left[\exp\left[-\beta\sum_{\tau} E(\tau)\boxed{\sum_{a=1}^{n}\delta\left(\mathbf{s}^a,\tau\right)}\right]\right]_{\mathbf{E}} = \exp\left[\sum_{t=1}^{n} p_t \frac{N(t\beta)^2}{4}\right]$$

replicas occupying τ

- Here, $(p_1, p_2, \dots, p_n)$ is a partition of n that satisfies

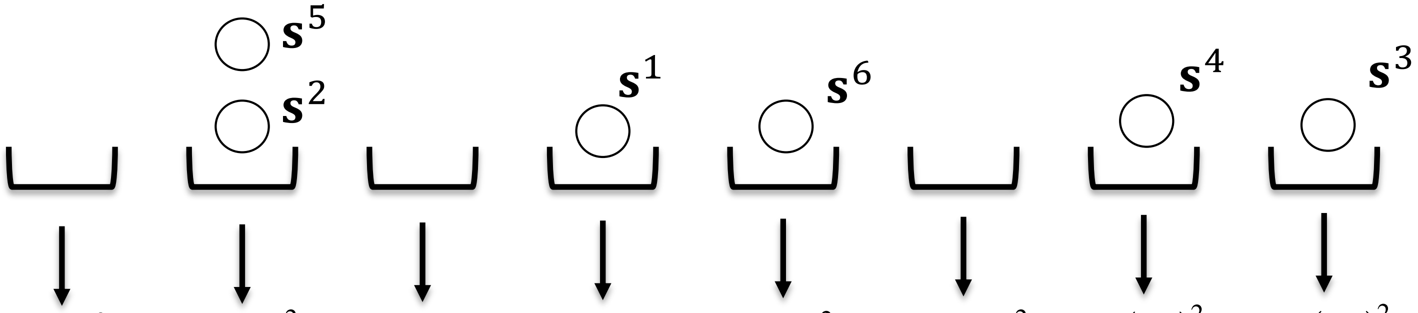
$$\begin{cases} p_t \geq 0 \ \ (t = 1,2,\dots,n) \\ p_1 + 2p_2 + \dots + np_n = n \end{cases}$$

  - Plays the role of "order parameter"

Partition of n and configuration

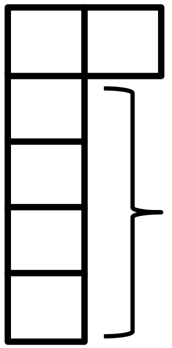
Ex) $N = 3 \to \tau \in \{+1,-1\}^3$ $n = 6$ replicas

State τ



$[\cdots]_{E(\tau)}$ $e^{\frac{N(0\beta)^2}{4}}$ $e^{\frac{N(2\beta)^2}{4}}$ $e^{\frac{N(0\beta)^2}{4}}$ $e^{\frac{N(1\beta)^2}{4}}$ $e^{\frac{N(1\beta)^2}{4}}$ $e^{\frac{N(0\beta)^2}{4}}$ $e^{\frac{N(1\beta)^2}{4}}$ $e^{\frac{N(1\beta)^2}{4}}$

$$(p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_9) = (4,1,0,0,0,0,0,0)$$



$p_2 = 1$

$p_1 = 4$

Expression by Young diagram

$$\left[\exp\left[-\beta \sum_{\tau\in\{+1,-1\}^2} E(\tau) \sum_{a=1}^{6} \delta(\mathbf{s}^a, \tau)\right]\right]_{\mathbf{E}} = \exp\left[1 \times \frac{N(2\beta)^2}{4} + 4 \times \frac{N(1\beta)^2}{4}\right]$$

$$= \exp\left[2N\beta^2\right]$$

Exact expression for $n = 1,2, \ldots$

- $[Z^n(\beta|\mathbf{E})]_{\mathrm{E}}$ is expressed exactly by a summation over partitions of n as

$$\left[Z^n \left(\beta | \mathbf{E} \right) \right]_{\mathbf{E}} = \sum_{(p_1, p_2, \ldots, p_n)} W\left(p_1, p_2, \ldots, p_n\right) \exp\left[\sum_{t=1}^{n} p_t \frac{N(t\beta)^2}{4} \right]$$

$W\left(p_1, p_2, \ldots, p_n\right)$: The number of microscopic configurations of replicas $\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^n$ that correspond to a partition of n, $(p_1, p_2, \ldots, p_n)$.

Also grows exponentially in N.

Concentration of measure

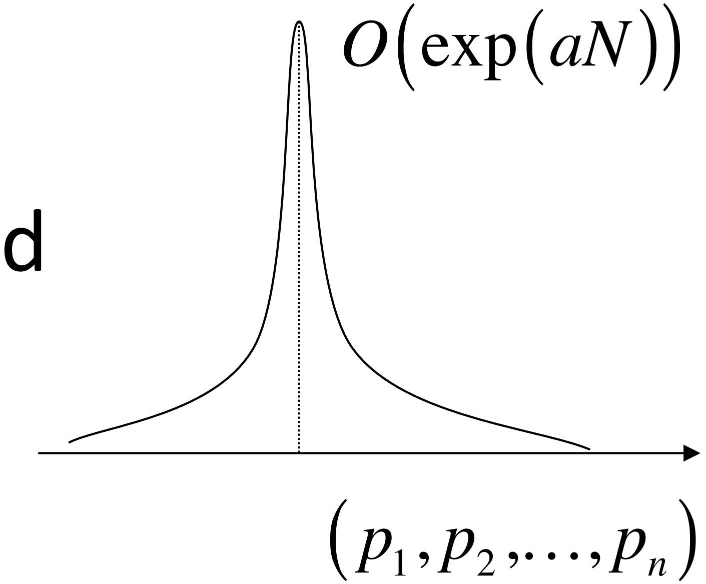
- The summation range of the partitions of n is finite independently of the system size N.
- On the other hand, each term grows exponentially w.r.t. N.

$$W\left(p_1,p_2,\ldots,p_n\right)\exp\left[\sum_{t=1}^{n} p_t \frac{N\left(t\beta\right)^2}{4}\right] \sim O\left(\exp\left(aN\right)\right)$$

$$N \to \infty$$

The moment can be represented by a single dominant term

(saddle point assessment)



Replica symmetry (RS) and RS ansatz

- To find the dominant term, we introduce the following assumption termed the replica symmetric (RS) ansatz

- **RS ansatz:**
  The expression of the moment
  $$\left[Z^n\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}} = \sum_{\mathbf{s}^1,\mathbf{s}^2,\ldots,\mathbf{s}^n}\left[\exp\left[-\beta\sum_{\tau}E(\tau)\sum_{a=1}^{n}\delta\left(\mathbf{s}^a,\tau\right)\right]\right]_{\mathbf{E}}$$
  is invariant under any permutation of replica indices $a = 1,2,\ldots,n$. This property is termed the *replica symmetry*. We assume that the dominant partition of n in the summation also satisfies this symmetry.

Two RS solutions

- Under RS ansatz, there are only two candidates of the dominant term

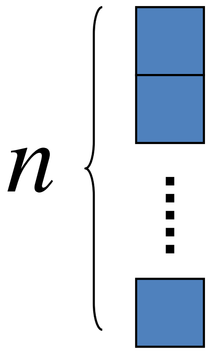
RS1

$$n=\underbrace{1+1+\ldots+1}_{n}$$

$$\Updownarrow$$

$$(p_1,p_2,\ldots,p_n)=(n,0,\ldots,0)$$

$$\Updownarrow$$



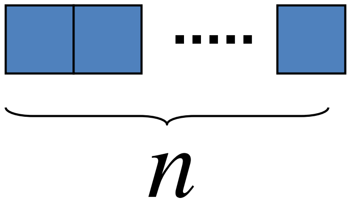
RS2

$$n=n$$

$$\Updownarrow$$

$$(p_1,p_2,\ldots,p_n)=(0,\ldots,0,1)$$

$$\Updownarrow$$



RS1

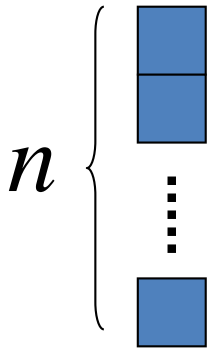
- $W(n, 0, \ldots, 0)$
  = #way of placing n replicas at n different states out of 2^N states
  $=2^N \times (2^N - 1) \times \cdots \times (2^N - n + 1) \simeq 2^{nN}$
- $\exp\left[n \times \frac{N\beta^2}{4}\right]$

$$\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} \simeq W(n,0,\ldots,0) \times \exp\left[n \times \frac{N\beta^2}{4}\right]$$

$$= \exp\left[Nn\left(\frac{\beta^2}{4} + \ln 2\right)\right]$$

$$(p_1, p_2, \ldots, p_n) = (n, 0, \ldots, 0)$$

$$\Updownarrow$$

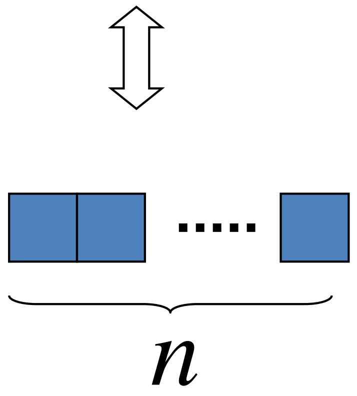


RS2

- $W(0,\ldots,0,1)$
  = #way of choosing a single state out of 2^N states at which all of the n replicas are placed
  =2^N

- $\exp\left[1\times\frac{N(n\beta)^2}{4}\right]$

$$(p_1, p_2, \ldots, p_n) = (0, \ldots, 0, 1)$$



$$\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} \simeq W(0,\ldots,0,1)\times\exp\left[1\times\frac{N(n\beta)^2}{4}\right]$$

$$= \exp\left[N\left(\frac{(n\beta)^2}{4} + \ln 2\right)\right]$$

Analytical continuation

- RS1

$$\left(\lim_{N\to\infty}\frac{1}{N}\ln\left[Z^n\left(\beta\middle|\mathbf{E}\right)\right]_{\mathbf{E}}=\right)\phi_{\mathrm{RS1}}\left(n,\beta\right)=n\left(\frac{\beta^2}{4}+\ln 2\right)$$

- RS2

$$\phi_{\mathrm{RS2}}\left(n,\beta\right)=\frac{\left(n\beta\right)^2}{4}+\ln 2$$

The both expressions can be defined for real numbers $n \in \mathbb{R}$

⇩

So, we analytically continue these expressions from $n = 1,2,\ldots$ to $n \in \mathbb{R}$, and use them for taking limit $n \to 0$.

Success/failure of the RS solutions

- RS1
  - Successfully reproduces the correct result for $\beta < \beta_c$

$$f(\beta) = -\frac{1}{\beta}\frac{\partial \phi_{\mathrm{RS1}}(n,\beta)}{\partial n} = -\frac{\beta}{4} - \frac{1}{\beta}\ln 2$$

- RS2
  - Leads to an obviously wrong answer

$$\lim_{n\to 0}\phi_{\mathrm{RS2}}(n,\beta) = \lim_{n\to 0}\left\{\frac{(n\beta)^2}{4} + \ln 2\right\} = \ln 2 \quad \text{yields} \quad \lim_{n\to 0}\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} = e^{N\ln 2} = 2^N$$

$$\neq 1$$

Low temperature behavior for $\beta > \beta_c$ cannot be reproduced ⇒ Limitation of the replica method?

Success/failure of the RS solutions

- RS1
  - Successfully reproduces the correct result for $\beta < \beta_c$

$$f(\beta) = -\frac{1}{\beta}\frac{\partial \phi_{\mathrm{RS1}}(n,\beta)}{\partial n} = -\frac{\beta}{4} - \frac{1}{\beta}\ln 2$$

- RS2
  - Leads to an obviously wrong answer

$$\lim_{n\to 0}\phi_{\mathrm{RS2}}(n,\beta) = \lim_{n\to 0}\left\{\frac{(n\beta)^2}{4} + \ln 2\right\} = \ln 2 \quad \text{yields} \quad \lim_{n\to 0}\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} = e^{N\ln 2} = 2^N$$

$\neq 1$

Lo be

re d?

Can't say for sure yet!

Still a possibility that the "RS ansatz" was wrong

One-step replica symmetry breaking (1RSB) solution

- Consider a candidate of *lower replica symmetry*
  - *Not fully symmetric. But, still, partially symmetric*

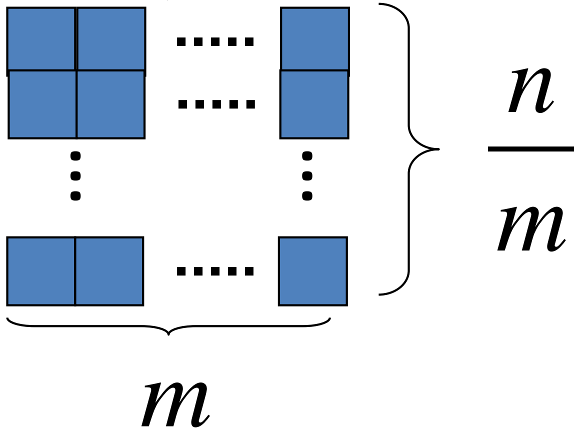
1RSB

$$n = \underbrace{m + m + \ldots + m}_{n/m}$$

$$\Updownarrow$$

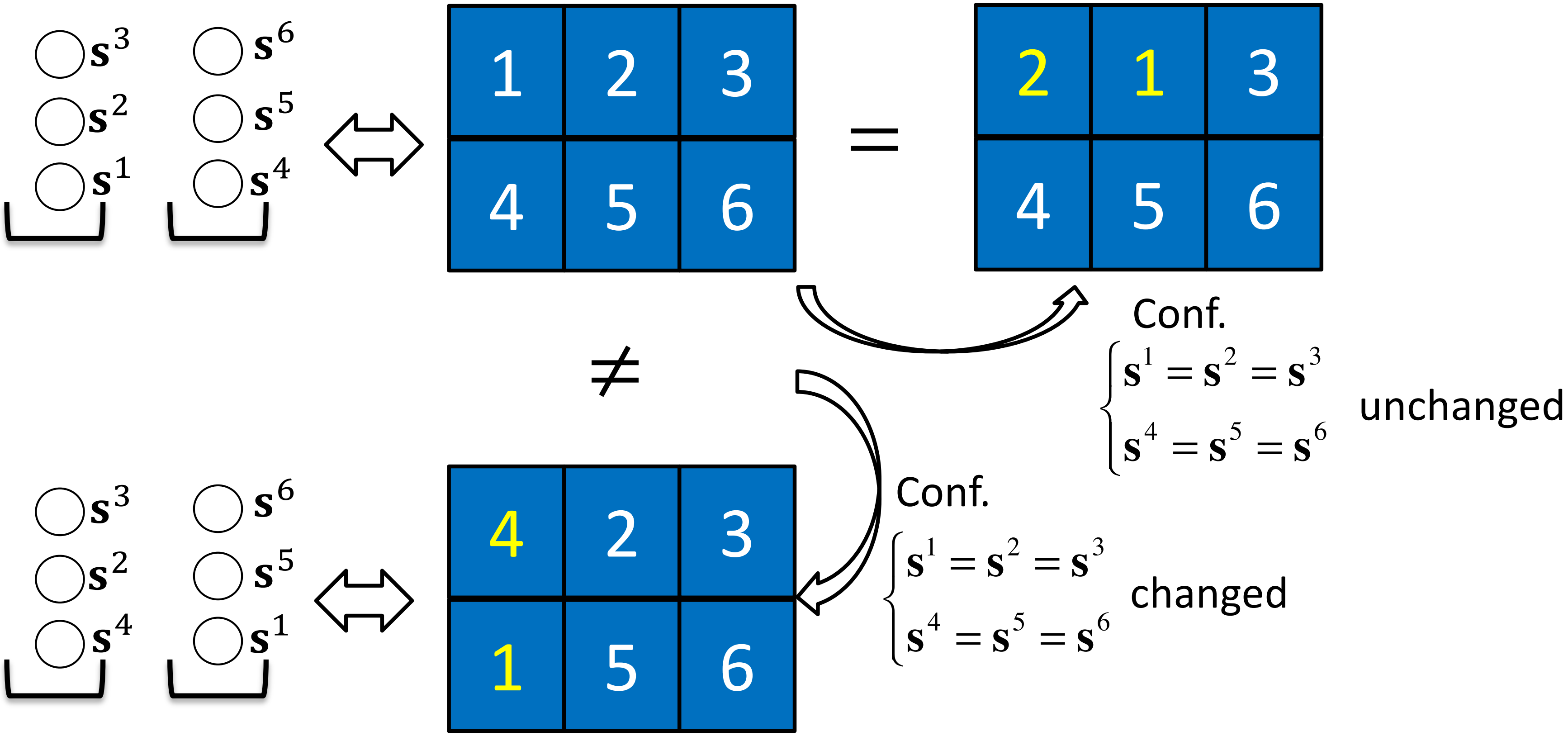
$$\left(p_1, p_2, \ldots, p_n\right) = \left(0, \ldots, 0, \frac{n}{m}, 0, \ldots, 0\right)$$

$$\Updownarrow$$



One-step replica symmetry breaking (1RSB) solution

- Consider a candidate of *lower replica symmetry*
  - *Not fully symmetric. But, still, partially symmetric*



1	2	3
4	5	6

=

2	1	3
4	5	6

≠

4	2	3
1	5	6

Conf.

$$\begin{cases} \mathbf{s}^1 = \mathbf{s}^2 = \mathbf{s}^3 \\ \mathbf{s}^4 = \mathbf{s}^5 = \mathbf{s}^6 \end{cases}$$

unchanged

Conf.

$$\begin{cases} \mathbf{s}^1 = \mathbf{s}^2 = \mathbf{s}^3 \\ \mathbf{s}^4 = \mathbf{s}^5 = \mathbf{s}^6 \end{cases}$$

changed

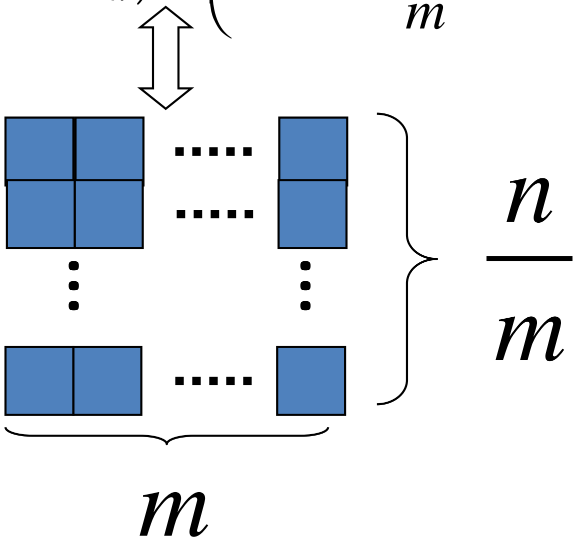
1RSB

- $W\left(0,\ldots,0,\frac{n}{m},0,\ldots,0\right)$
  = #way of choosing $\frac{n}{m}$ states out of 2^N states and distributing n replicas to them by equal size m
  $=2^N\times\cdots\times\left(2^N-\frac{n}{m}+1\right)\times\frac{n!}{(m!)^{\frac{n}{m}}}\simeq 2^{Nn/m}$
- $\exp\left[\frac{n}{m}\times\frac{N(m\beta)^2}{4}\right]$

$$\left(p_1,p_2,\ldots,p_n\right)=\left(0,\ldots,0,\frac{n}{m},0,\ldots,0\right)$$

$$\left[Z^n\left(\beta|\mathbf{E}\right)\right]_{\mathbf{E}}\simeq W\left(0,\ldots,0,\frac{n}{m},0,\ldots,0\right)$$

$$\times\exp\left[\frac{n}{m}\times\frac{N\left(m\beta\right)^2}{4}\right]=\exp\left[\frac{nN}{m}\left(\frac{\left(m\beta\right)^2}{4}+\ln 2\right)\right]$$



Analytical continuation

- 1RSB
  - We determine the breaking parameter m by extremization

$$\phi_{\mathrm{1RSB}}(n,\beta)=\operatorname*{extr}_{m}\left\{\frac{n}{m}\left(\frac{(m\beta)^2}{4}+\ln 2\right)\right\}$$

$$=n\beta\sqrt{\ln 2}\qquad\left(m^*(\beta)=2\sqrt{\ln 2}/\beta\right)$$

  - This successfully reproduces the correct low temperature solution for $\beta > \beta_c = 2\sqrt{\ln 2}$ as

$$f(\beta)=-\frac{1}{\beta}\lim_{n\to 0}\frac{\partial\phi_{\mathrm{1RSB}}(n,\beta)}{\partial n}=-\sqrt{\ln 2}$$

In the end…

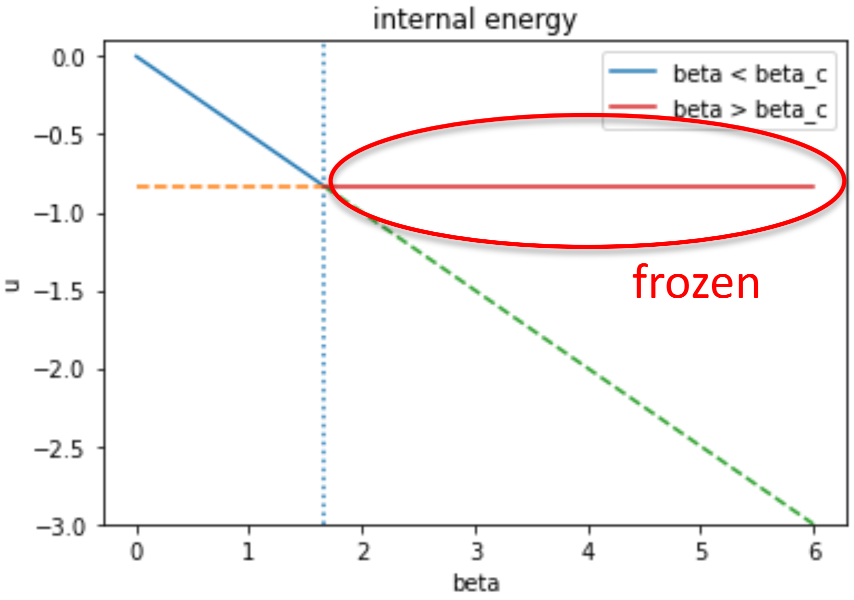
- Taking into account RSB reproduced the correct answer for REM

$$f(\beta)=\begin{cases} -\frac{\beta}{4}-\frac{\ln 2}{\beta}, & \beta<\beta_c\left(=2\sqrt{\ln 2}\right) \qquad \text{(RS1)} \\ -\sqrt{\ln 2}, & \beta>\beta_c \qquad \text{(1RSB)} \end{cases}$$

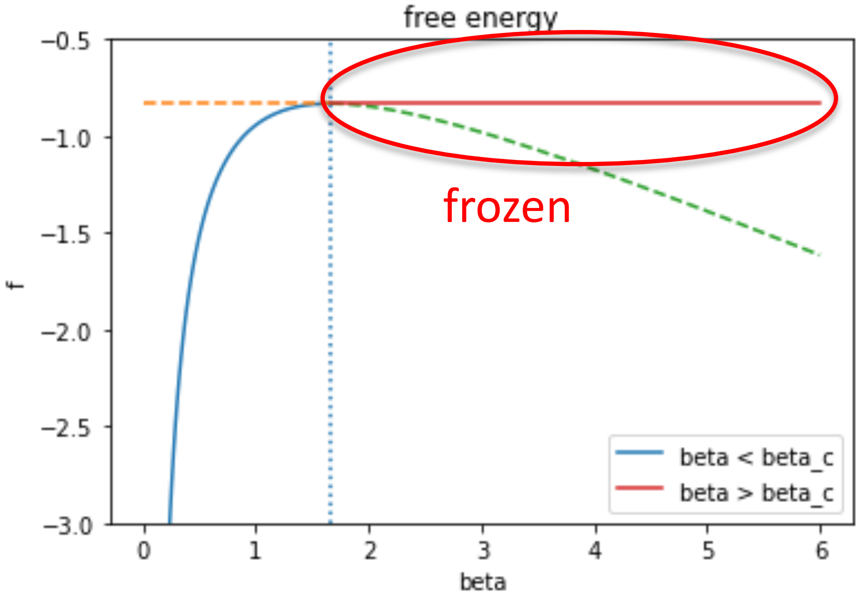
$$u(\beta)=\frac{\partial\left(\beta f(\beta)\right)}{\partial\beta}=\begin{cases} -\frac{\beta}{2}, & \beta<\beta_c \\ -\sqrt{\ln 2}, & \beta>\beta_c \end{cases}$$

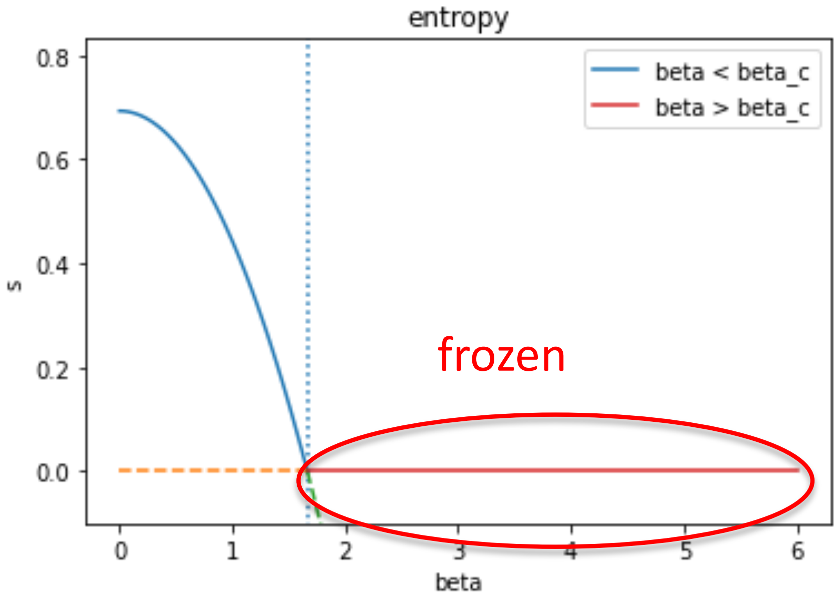
$$s(\beta)=\beta\left(u(\beta)-f(\beta)\right)=\begin{cases} -\frac{\beta^2}{4}+\ln 2, & \beta<\beta_c \\ 0, & \beta>\beta_c \end{cases}$$

Phase transition



“Frozen” for $\beta > \beta_c$





Discussion

- However, the argument is somewhat *ad hoc* and looks little principled

Is there any guideline or bird-eye's view behind the calculation?

Perspective of large deviation statistics

- Replica method = a technique to evaluate not a single value but a ***distribution*** of the value of $\ln Z(\beta|\mathbf{E})$
- The value of $\ln Z(\beta|\mathbf{E})$ is $O(N)$, and therefore, is expected to obey ***large deviation statistics***

Large deviation statistics

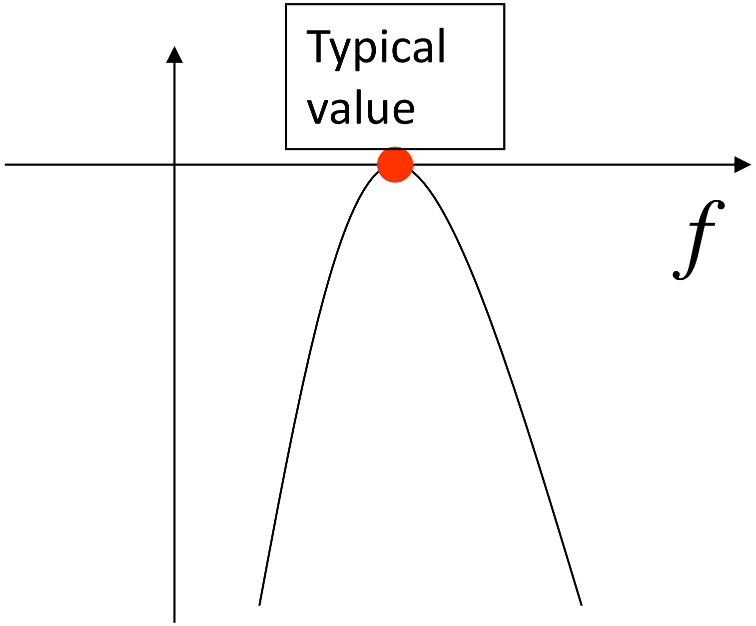
$$Z(\beta|\mathbf{E}) \sim e^{-N\beta f}$$

$$P(f|\beta) \sim e^{Nc(f,\beta)}$$

$$\underline{\left(c(f,\beta)\le 0\right)}$$

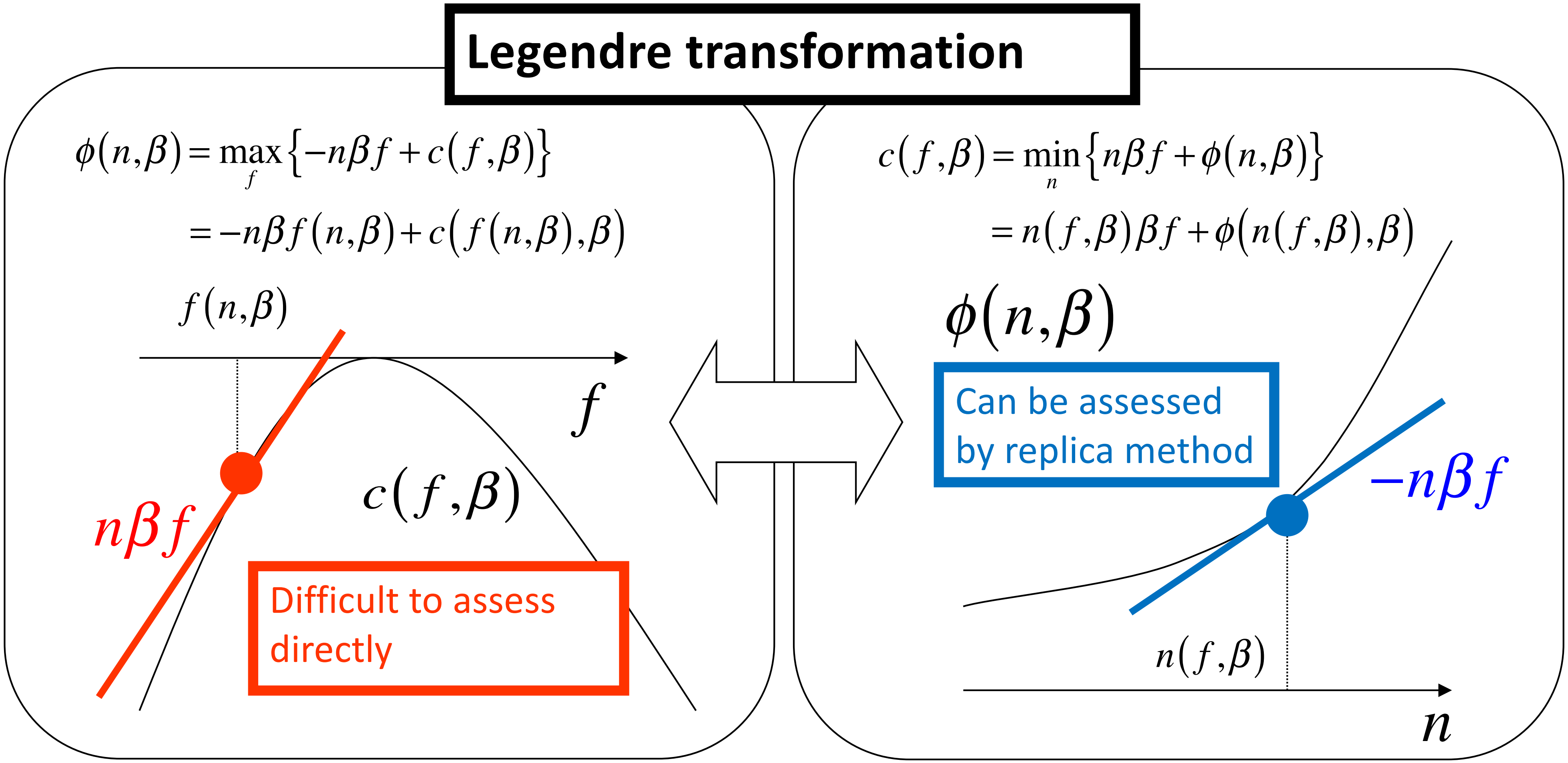
Rate function

$$c(f,\beta)=\frac{1}{N}\ln P(f|\beta)$$



$$\left[Z^n(\beta|\mathbf{E})\right]_{\mathbf{E}} \triangleq \exp\left[\underline{N\phi(n,\beta)}\right] = \int df \underset{\overset{\Uparrow}{Z^n}}{e^{-Nn\beta f}} \times \underset{\overset{\Uparrow}{P(Z)}}{e^{Nc(f,\beta)}}$$

$$\sim \exp\left[\underline{N\max_{f}\left\{-n\beta f + c(f,\beta)\right\}}\right]$$



Legendre transformation

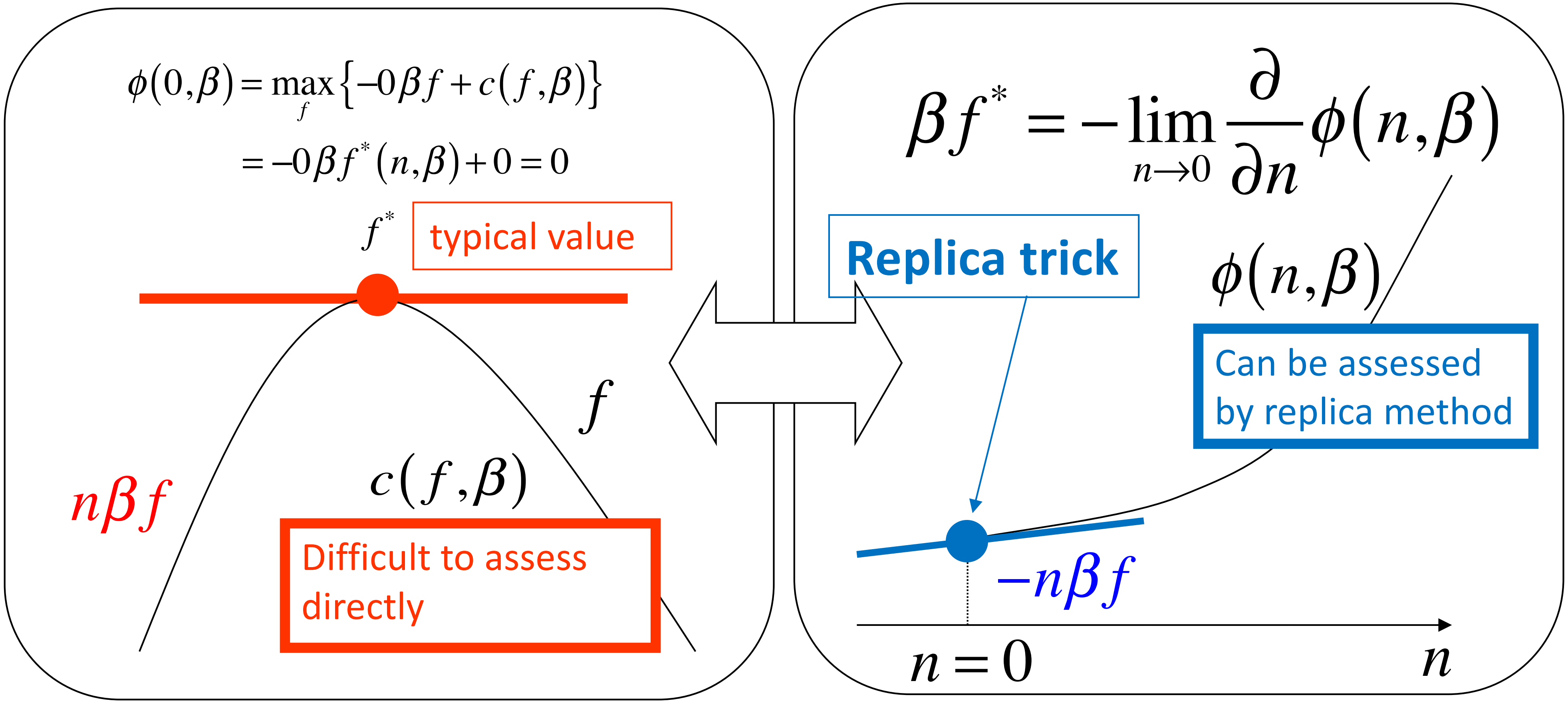
$$\phi(n,\beta) = \max_{f}\left\{-n\beta f + c(f,\beta)\right\}$$
$$= -n\beta f(n,\beta) + c\left(f(n,\beta),\beta\right)$$

Difficult to assess directly

$$c(f,\beta) = \min_{n}\left\{n\beta f + \phi(n,\beta)\right\}$$
$$= n(f,\beta)\beta f + \phi\left(n(f,\beta),\beta\right)$$

Can be assessed by replica method

- $\phi(n,\beta) \Longleftrightarrow c(f,\beta)$
  - $n \to 0$ corresponds to the typical (highest prob.) case
  - Atypical cases can be analyzed by finite n as well



$$\phi(0,\beta) = \max_f\left\{-0\beta f + c(f,\beta)\right\}$$

$$= -0\beta f^*(n,\beta) + 0 = 0$$

f^* typical value

f

$c(f,\beta)$

$n\beta f$

Difficult to assess directly

$$\beta f^* = -\lim_{n\to 0}\frac{\partial}{\partial n}\phi(n,\beta)$$

Replica trick

$\phi(n,\beta)$

Can be assessed by replica method

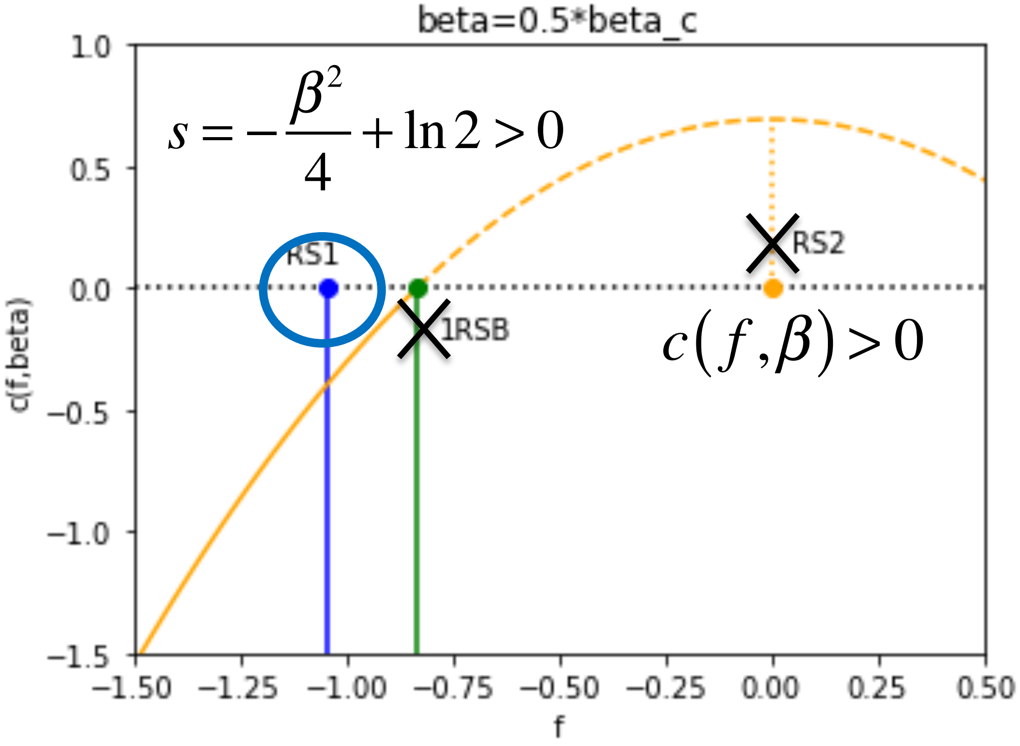
$-n\beta f$

$n = 0$

n

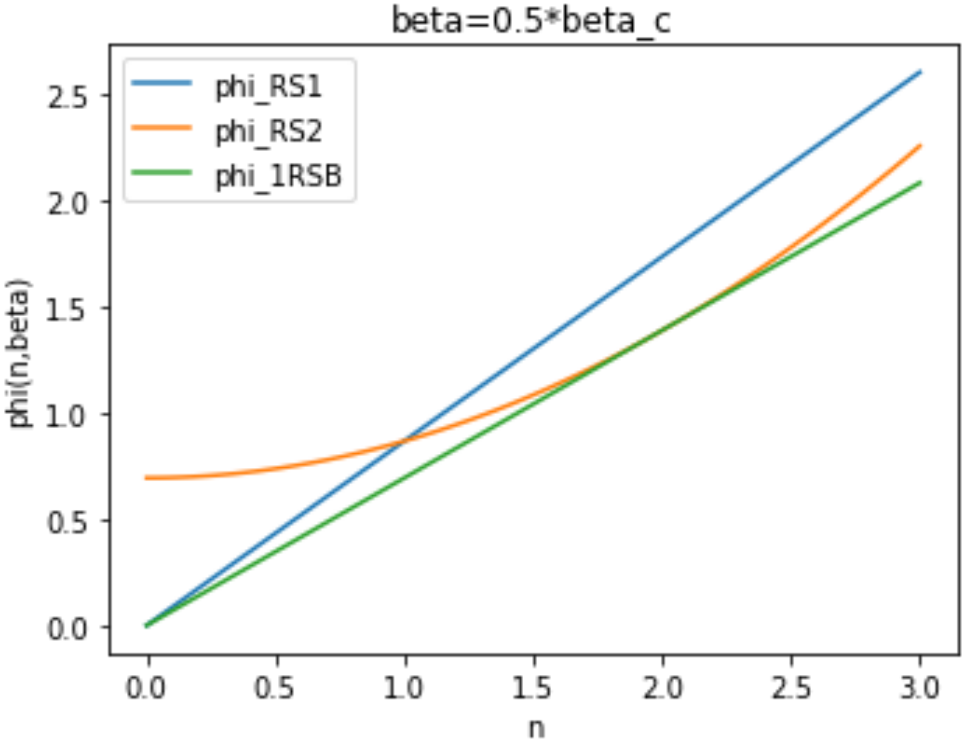
Selection of appropriate solution

$$\beta < \beta_c$$



$s = -\frac{\beta^2}{4} + \ln 2 > 0$

$c(f, \beta) > 0$



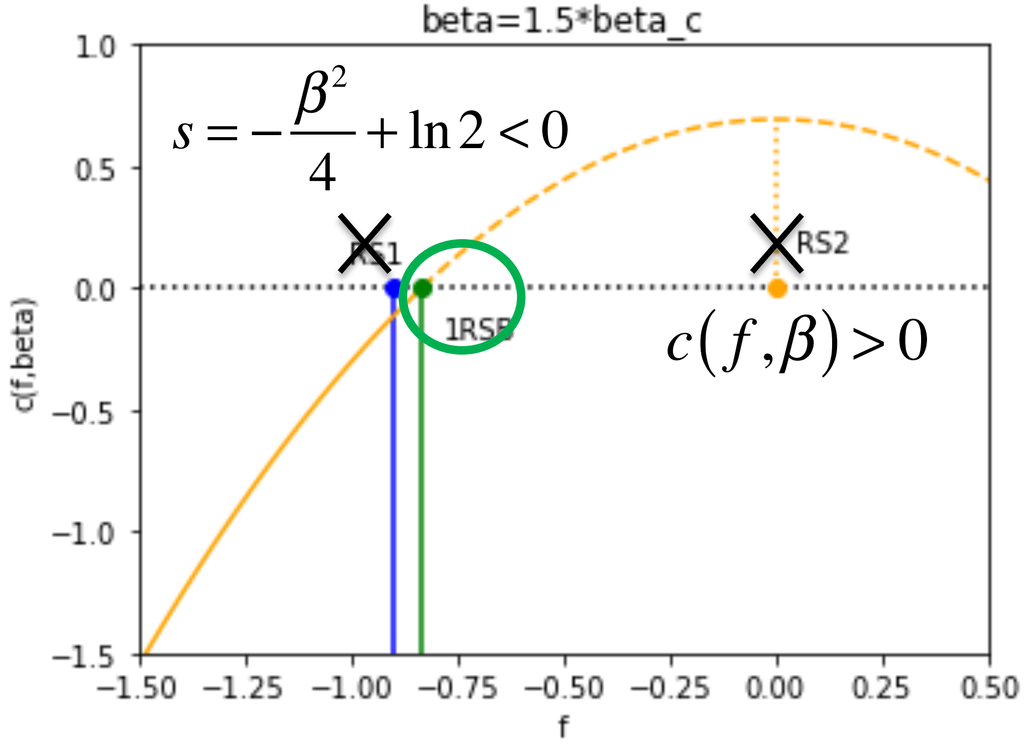
RS1: appropriate

RS2: inappropriate as rate function is positive

1RSB: inappropriate as f_{1RSB} is higher than f_{RS1}

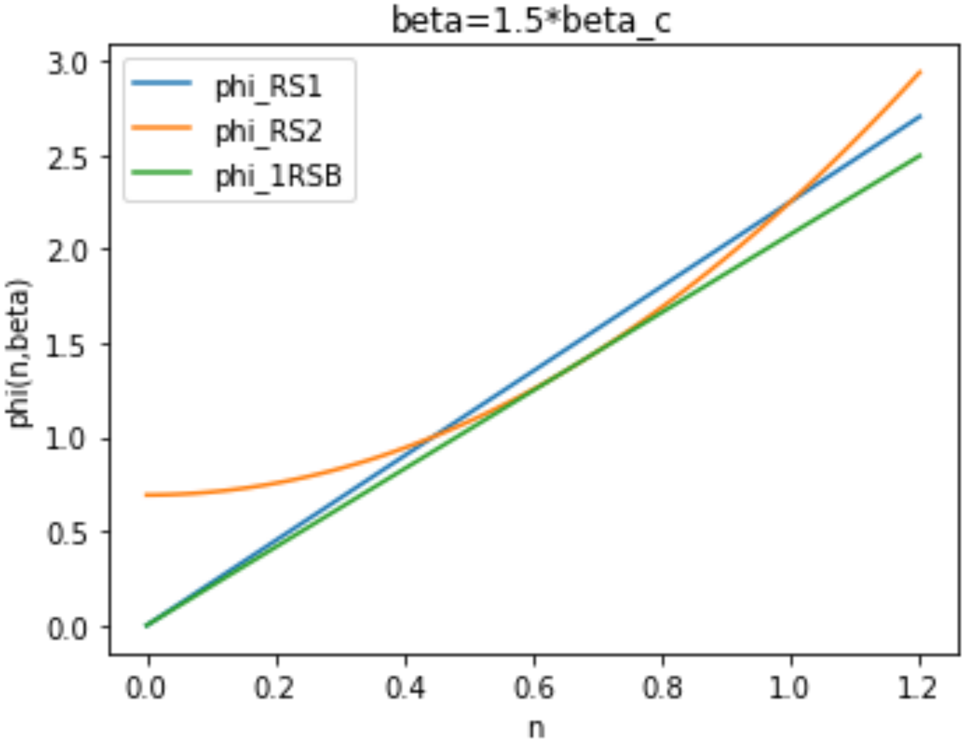
Selection of appropriate solution

$$\beta > \beta_c$$



$$s = -\frac{\beta^2}{4} + \ln 2 < 0$$

$$c(f, \beta) > 0$$



RS1: inappropriate as entropy is negative

RS2: inappropriate as rate function is positive

1RSB: appropriate

Large deviation perspective of 1RSB solution

Dist. of the lowest energy in REM

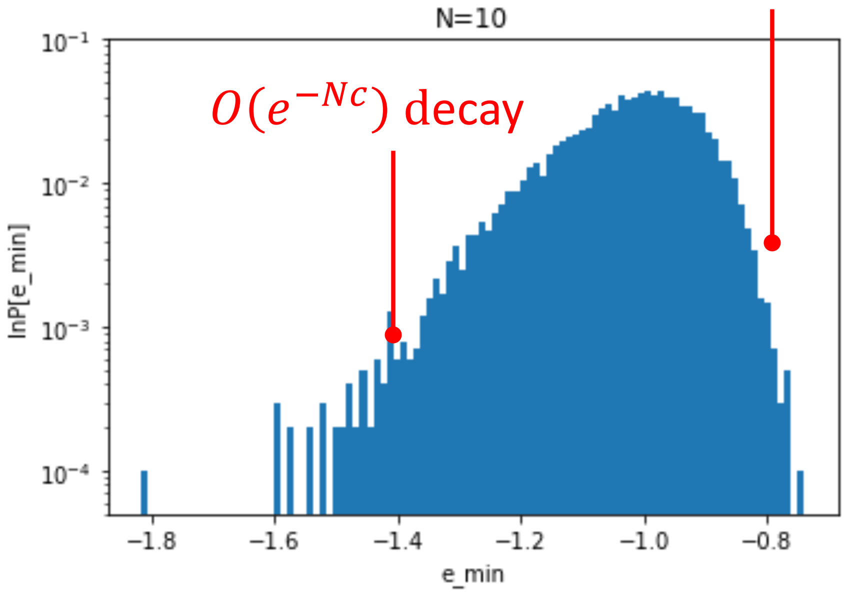
$$P\left[\min_{\tau}\left\{E(\tau)\right\}=Ne_{\min}\right]$$

$$=P\left[E(\tau)=Ne_{\min}\ \text{for}\ \exists\tau\ \text{and}\ E(\tau')>Ne_{\min}\ \text{for the other states}\ \tau'\right]$$

$$=2^{N}\times P\left(E=Ne_{\min}\right)\times\left(\int_{Ne_{\min}}^{+\infty}P(E)dE\right)^{2^{N}-1}$$

$$\simeq\begin{cases}\exp\left(N\left(\ln 2-e_{\min}^{2}\right)\right), & e_{\min}<-\sqrt{\ln 2}\\ \exp(-\exp\left(N\left(\ln 2-e_{\min}^{2}\right)\right)), & e_{\min}>-\sqrt{\ln 2}\end{cases}$$

(Gumbel distribution)



$O\left(e^{-Nc}\right)$ decay

$O\left(e^{-e^{Nc}}\right)$ decay

Large deviation perspective of 1RSB solution

For $N \gg 1$,

$$f = -\frac{1}{N\beta}\ln\left[\sum_s \exp\left[-\sum_\tau \beta E(\tau)\delta(s,\tau)\right]\right] \simeq \frac{\min_\tau E(\tau)}{N} = e_{\min}$$

holds for $\beta > \beta_c$. This implies the rate function for $\beta > \beta_c$ is given as

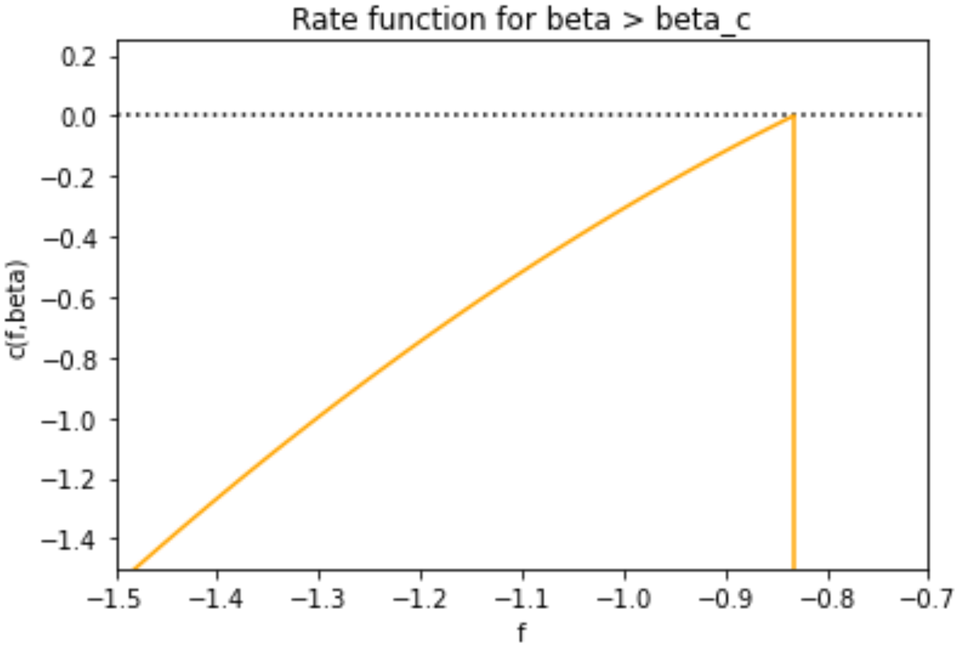
$$c(f,\beta) \simeq \begin{cases} \ln 2 - f^2, & f < -\sqrt{\ln 2} \\ -\infty, & f > -\sqrt{\ln 2} \end{cases}$$

independently of β.

⇩

Satisfies constraint

$$c(f,\beta) \le 0$$



Large deviation perspective of 1RSB solution

1RSB corresponds to the procedure to incorporate the constraint $c(f,\beta) \le 0$ for the RS2 solution.

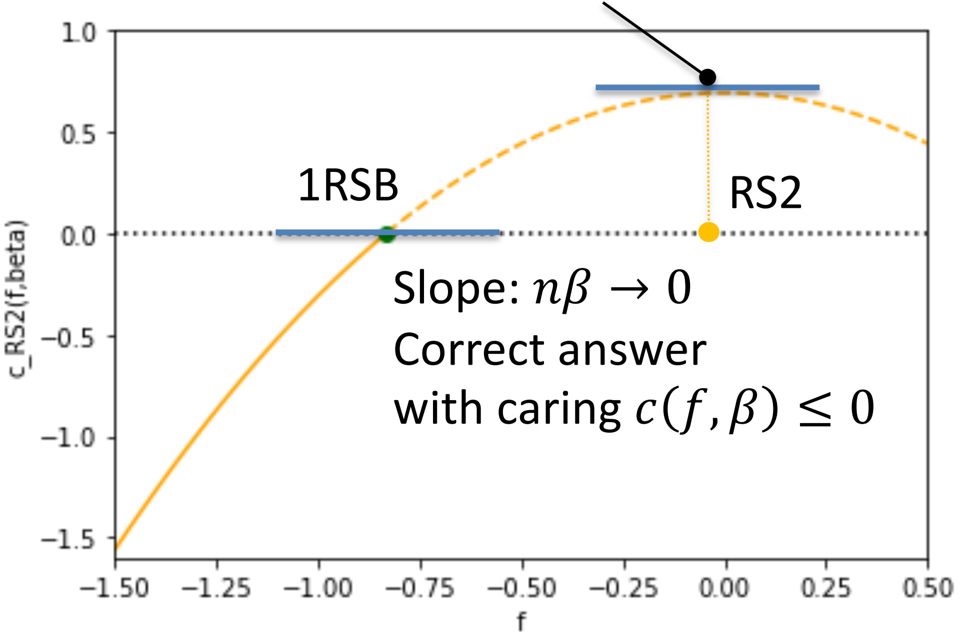
$$\phi_{\mathrm{RS2}}(n,\beta)=\max_{f}\left\{-n\beta f+c_{\mathrm{RS2}}(f,\beta)\right\}$$

$$\to \max_{f,c_{\mathrm{RS2}}(f,\beta)\le 0}\left\{-n\beta f+c_{\mathrm{RS2}}(f,\beta)\right\}$$

$$=\max_{f,\ln 2-f^2\le 0}\left\{-n\beta f+\ln 2-f^2\right\}$$

$$=\begin{cases}\dfrac{(n\beta)^2}{4}+\ln 2 & \left(n\ge 2\sqrt{\ln 2}/\beta\right)\\ \underline{n\beta\sqrt{\ln 2}} & \left(n< 2\sqrt{\ln 2}/\beta\right)\end{cases}$$

Wrong answer
without caring $c(f,\beta)\le 0$



Slope: $n\beta\to 0$
Correct answer
with caring $c(f,\beta)\le 0$

$$\phi_{\mathrm{1RSB}}(n,\beta)=\operatorname*{extr}_{m}\left\{\frac{n}{m}\left(\frac{(m\beta)^2}{4}+\ln 2\right)\right\}=\underline{n\beta\sqrt{\ln 2}}$$

Summary for 2nd part

- Demonstrated the computation of the replica method (RM) for random energy model (REM)
  - A testbed for RM as exactly solvable without using RM

- The computation indicated
  - Exact result is reproduced for $\beta < \beta_c$ by the saddle point assessment under the replica symmetric (RS) ansatz
  - Meanwhile, RS ansatz does not lead to the correct result for $\beta > \beta_c$. This implies that the RS ansatz is inappropriate for $\beta > \beta_c$.
  - Therefore, we introduced an assumption of lower replica symmetry, the 1-step replica symmetry breaking ansatz (1RSB), which reproduces the correct result for $\beta > \beta_c$

  - Consideration based on large deviation statistics shows that the emergence of the 1RSB solution originates from the singularity of the distribution of the lowest energy value (Gumbel dist.) in REM

Discussion for 2nd part

- The calculation illustrates general recipe of the replica method as follows:
  1. Construct a solution under the RS ansatz
  2. Check if it is mathematically consistent
     - Positivity of entropy
     - Negativity of rate function
     - Stability of saddle point
     - …
  3. If all the all check points are passed, keep it as a “tentative candidate” of the correct solution
  4. Otherwise, return to 1. using a certain RSB ansatz if necessary, until a mathematically consistent solution is found

Discussion for 2nd part

- Applicability to other problems:
  - One can handle random coding ensemble (RCE) in a similar manner, which provides an equivalent result with Shannon's channel coding theorem

  - On the other hand, phase transitions of other types occur for random k-SAT problems, and slightly different treatment is necessary.
    However, construction of solutions under one (or more, if necessary) step replica symmetry breaking (RSB) ansatz still provides the correct results after the phase transitions.

  - Clarification of the reason why appropriate RSB schemes generally provide the correct results even after the phase transitions is an open problem

Discussion for 2nd part

- About the cavity method:
  - Another physics-based technique usable as efficient inference/optimization algorithms
  - Generalization of Bethe (tree) approximation applicable to disordered systems defined over graphs (not applicable to REM)

Joint dist.

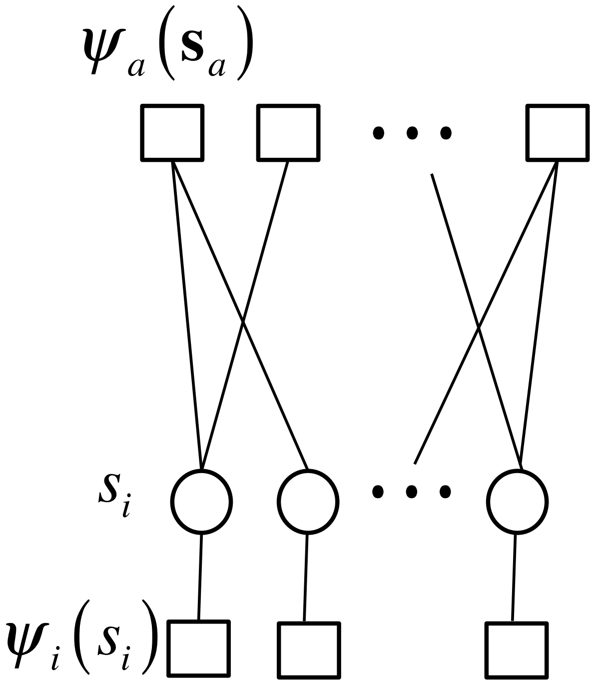
$$P(\mathbf{s}) \propto \prod_a \psi_a(\mathbf{s}_a) \prod_i \psi_i(s_i)$$

Belief propagation

$$\begin{cases} m_{a\to i}(s_i) = \alpha_{a\to i} \sum_{\mathbf{s}_a \backslash s_i} \psi_a(\mathbf{s}_a) \prod_{j\in\partial a\backslash i} m_{j\to a}(s_j) \\ m_{i\to a}(s_i) = \alpha_{i\to a} \psi_i(s_i) \prod_{b\in\partial i\backslash a} m_{b\to i}(s_i) \end{cases}$$

Marginal dist.

$$P(s_i) = \sum_{\mathbf{s}\backslash s_i} P(\mathbf{s}) \simeq \alpha_i \psi_i(s_i) \prod_{a\in\partial i} m_{a\to i}(s_i)$$



Thank you for your listening