Exact Analytic Continuation with Respect to the Replica Number in the Discrete Random Energy Model of Finite System Size

Kenzo OGURE[1,*] and Yoshiyuki KABASHIMA[2,**]

[1] *Theory Group, Institute for Cosmic Ray Research, Kashiwa 277-8582, Japan*
[2] *Department of Computational Intelligence and Systems Science,*
Tokyo Institute of Technology, Yokohama 226-8502, Japan

An expression for the moment of the partition function valid for any finite system size N and complex power n [$\mathrm{Re}(n) > 0$] is obtained for a simple spin glass model termed the *discrete random energy model* (DREM). We investigate the behavior of this moment in the thermodynamic limit, $N \to \infty$, using this expression, and we find that a phase transition occurs at a certain real value of the replica number when the temperature is sufficiently low. This represents a direct clarification of the scenario of replica symmetry breaking of the DREM in the replica number space *without use of the replica trick*. The validity of the expression is confirmed numerically.

§1. Introduction

The replica method (RM) is one of the few analytic schemes available to the study of disordered systems.[1] In physics, this method has been well known since the 1970s and has been successfully applied to the analysis of spin-glass models,[2]–[4] although the essential idea behind the method can be dated back to the end of 1920s, when it appeared as a theorem for computing the average of logarithms.[5]–[7] More recently, considerable attention has been paid to the similarity between the statistical mechanics of disordered systems and the Bayesian method in problems related to information processing (IP).[8] Accordingly, the number and variety of applications of the RM to problems in IP research are increasing rapidly. They include error-correcting codes,[9],[10] image restoration,[11],[12] neural networks,[13] combinatorial problems,[14],[15] and so on.

Although only the limiting value $(1/N) \langle \ln Z \rangle = \lim_{n \to 0} \left(\langle Z^n \rangle^{1/N} - 1 \right)/n$ is usually considered relevant, the RM can be considered a systematic procedure for calculating generalized moments $\langle Z^n \rangle$ of the partition function Z in the case $N \to \infty$ when Z depends on a certain external randomness. Here, N characterizes the system size, $n \in \mathbf{R}$(or $\mathbf{C}$) is a real (or complex[16]) number, and $\langle \cdots \rangle$ represents the average over the external randomness. For most problems, a direct assessment of $\langle Z^n \rangle$ is difficult for general $n \in \mathbf{R}$(or $\mathbf{C}$), although such an assessment for natural numbers, $n = 1, 2, \cdots$, is possible in the thermodynamic limit $N \to \infty$. Therefore, $\langle Z^n \rangle$ is first computed for natural numbers, and then their analytic continuation is used to extend $\langle Z^n \rangle$ to $n \in \mathbf{R}$(or $\mathbf{C}$). This is usually termed the *replica trick*.

*) E-mail: ogure@icrr.u-tokyo.ac.jp
**) E-mail: kaba@dis.titech.ac.jp

However, the validity of the replica trick is doubtful. The most obvious analytic continuation, obtained under the replica symmetric (RS) ansatz, sometimes leads to incorrect results. The causes of these errors were actively debated in the 1970s, until Parisi developed the replica-symmetry-breaking (RSB) scheme for constructing reasonable solutions within the framework of the RM.[4] Since this development, no examples have been found for which physically incorrect results have been derived using the RM, in conjunction with the Parisi scheme if necessary. Therefore, the RM is now empirically recognized as a reliable procedure in physics, although the mathematical justification of the replica trick remains an open question. However, this problem is now generating interest again, in particular, in the application of the RM to IP problems. This is because most theories in IP research have conventionally been developed with a high regard for mathematical soundness.[17,18]

The purpose of this paper is to provide a method to approach the problems of the RM. Specifically, we give a useful formula to compute $\langle Z^n \rangle$ *directly* for $n \in \boldsymbol{C}$ at *finite N* for a simple spin glass model, termed the discrete random energy model (DREM).[19]–[21] This formula is numerically tractable, and therefore one can directly observe how the system approaches the thermodynamic limit with the aid of a numerical calculation. Furthermore, this formula *analytically* clarifies the correct behavior for $N \to \infty$, making a direct examination of the validity of the RM.

We have two main reasons for picking the DREM from among the various disordered systems. First, this model is simple enough to handle analytically. It is already known that the RM, in conjunction with the Parisi scheme, allows for the correct evaluation of the free energy for a family of random energy models (REMs), including the DREM, in the limit of $n \to 0$.[19] However, the existing procedure seems at odds with a theorem concerning analytic continuation proven by Carlson,[22,23] which holds for the DREM of finite N, asserting the uniqueness of analytic continuation from natural numbers $n \in \boldsymbol{N}$ to complex numbers $n \in \boldsymbol{C}$, when the temperature is sufficiently low. For this, our approach shows that a phase transition occurs at a certain critical value if the replica number, $n_c \in [0,1]$, in such cases. This demonstrates that the RM can be consistent with Carlson's theorem. This may provide a useful guideline to carry out analytic continuation from $n \in \boldsymbol{N}$ to $n \in \boldsymbol{R}$ (or $\boldsymbol{C}$) in the RM. The second reason for piking the DREM is that there exits a relationship between the REM and certain problems of IP. Recent research on error-correcting codes has revealed that the REM is closely related to a randomly constructed code.[9,10] Such codes are known to provide the best error correction performance in information theory,[25] and the performance evaluation of such codes is similar to the computation of $\langle Z^n \rangle$ for $n \in \boldsymbol{R}$[26] (see Appendix A). Therefore, the current investigation should indirectly verify the RM-based analysis of error-correcting codes performed previously.[10,27,28]

This paper is organized as follows. In §2 we introduce the DREM and briefly review how the RM has been employed in the conventional analysis of this system. Comparing with Carlson's theorem, we address why the conventional scenario for taking a limit $n \to 0$ seems suspect. In order to resolve this difficulty, we propose in §3 a new scheme to directly evaluate $\langle Z^n \rangle$ for the REM with finite N and complex n without using the replica trick. Taking the limit $N \to \infty$, we analytically clarify

how $\langle Z^n \rangle$ behaves in the thermodynamic limit and numerically verify this behavior. In §4, we show how the RM can be consistent with the results obtained using the proposed scheme. Section 5 consists of a summary.

§2. The replica method in the discrete random energy model (DREM)

In order to clearly state the problem addressed in this paper, we first review how the RM has been conventionally employed in analyzing the REM.[19),29)] For convenience in the later analysis, we mainly concentrate on the DREM, but the problem addressed here is shared widely with other versions of the REM as well.

A DREM is composed of 2^N states, whose energies, ϵ_A ($A = 1, 2, \cdots, 2^N$), are independently drawn from an identical distribution,

$$P(E_i) = 2^{-M} \left(\begin{array}{c} M \\ \frac{1}{2}M + E_i \end{array} \right), \quad \text{with} \quad \left(E_i = i - \frac{M}{2} \right), \tag{2·1}$$

over $M + 1$ energy levels, $E_i = -M/2, -M/2 + 1, \cdots, M/2 - 1, M/2$. For each realization $\{\epsilon_A\}$, the partition function

$$Z = \sum_{A=1}^{2^N} \exp(-\beta \epsilon_A) \tag{2·2}$$

and the free energy (density)

$$F = -\frac{kT}{N} \log Z \tag{2·3}$$

can be used for computing various thermal averages. However, when the *configurational average* is required, one has to compute the average free energy, $\langle F \rangle = -\frac{kT}{N} \langle \log Z \rangle$, whose direct evaluation is generally difficult. Here, $\langle \cdots \rangle$ represents the configurational average with respect to $\{\epsilon_A\}$. On the other hand, the moments of the partition function $\langle Z^n \rangle$ can be easily calculated in various models for the natural numbers, $n = 1, 2, \cdots$. Therefore, the replica method evaluates the average free energy using the *replica trick*,

$$\frac{1}{N} \langle \log Z \rangle = \lim_{n \to 0} \frac{\langle Z^n \rangle^{\frac{1}{N}} - 1}{n}, \tag{2·4}$$

analytically continuing the expression of $\langle Z^n \rangle$ for $n = 1, 2, \cdots$ to that for real (or complex) numbers n.

For a given natural number n, the moment of the DREM is calculated as

$$\langle Z^n \rangle = \sum_{A_1=1}^{2^N} \sum_{A_2=1}^{2^N} \cdots \sum_{A_n=1}^{2^N} \sum_{i_1=0}^{M} P\left(E_{i_1}^{(1)} \right) \sum_{i_2=0}^{M} P\left(E_{i_2}^{(2)} \right) \cdots \sum_{i_{2^N}=0}^{M} P\left(E_{i_{2^N}}^{(2^N)} \right)$$

$$\exp\left(-\beta \sum_{B=1}^{2^N} E_{i_B}^{(B)} \sum_{\mu=1}^{n} \delta_{BA_\mu} \right)$$

$$= e^{\frac{nM\beta}{2}} \sum_{A_1=1}^{2^N} \sum_{A_2=1}^{2^N} \cdots \sum_{A_n=1}^{2^N} \prod_{B=1}^{2^N} \left(\frac{1 + e^{-\beta \sum_{\mu=1}^{n} \delta_{BA_\mu}}}{2} \right)^M$$

$$= \sum_{A_1=1}^{2^N} \sum_{A_2=1}^{2^N} \cdots \sum_{A_n=1}^{2^N} \prod_{B=1}^{2^N} \exp\left(N\alpha I \left(\beta \sum_{\mu=1}^{n} \delta_{BA_\mu} \right) \right) \tag{2.5}$$

where $\alpha = M/N$, and the function $I(x)$ is defined as

$$I(x) = \log\left(\cosh\frac{x}{2} \right). \tag{2.6}$$

The identity $\prod_{B=1}^{2^N} \exp\left(-\frac{\beta}{2} \sum_{\mu=1}^{n} \delta_{BA_\mu} \right) = \exp\left(-\frac{n\beta}{2} \right)$ was employed to obtain the final expression of Eq. (2.5). From this point, we focus on the case $\alpha > 1$, in which the replica symmetry can be broken when the temperature is sufficiently low.

Unfortunately, performing the summation in Eq. (2.5) exactly is difficult. Instead, in the conventional RM, Eq. (2.5) is represented by the most dominant contribution in the summation. This can be justified for natural numbers n in the limit $N \to \infty$. Notice that the summation is invariant with respect to the permutation of the replica indices $\mu = 1, 2, \cdots, n$. This *replica symmetry* narrows the set of candidates for the most dominant contribution to three possibilities, which are here referred to as solutions of replica symmetric 1 (RS1), replica symmetric 2 (RS2), and 1-step replica symmetry breaking (1RSB).

- RS1
  In RS1, all n replicas are assumed to occupy n different states $B (= 1, 2, \cdots, 2^N)$. Therefore, for a given B, we have

$$\sum_{\mu=1}^{n} \delta_{BA_\mu} = \left\{ \begin{array}{ll} 1 & \text{when } B \text{ is one of the } n \text{ occupied states,} \\ 0 & \text{otherwise.} \end{array} \right. \tag{2.7}$$

The number of ways to assign n replicas to n out of 2^N different states is

$$2^N \times (2^N - 1) \times \cdots \times (2^N - n + 1) \sim 2^{nN}. \tag{2.8}$$

For each case, the configuration of this type contributes an amount

$$\exp\left(nN\alpha I(\beta) \right) \tag{2.9}$$

in Eq. (2.5). This means that the contribution to the moment from RS1 becomes

$$\langle Z^n \rangle = 2^{nN} \exp\left(nN\alpha I(\beta) \right)$$
$$= \exp nN \left(\log 2 + \alpha \log\left(\cosh\frac{\beta}{2} \right) \right). \tag{2.10}$$

- RS2
  In RS2, all n replicas are assumed to occupy a particular state B. Therefore, for a given B, we have

$$\sum_{\mu=1}^{n} \delta_{BA_\mu} = \left\{ \begin{array}{ll} n & \text{when } B \text{ is the occupied state,} \\ 0 & \text{otherwise.} \end{array} \right. \tag{2.11}$$

The number of ways to choose one out of 2^N states is 2^N. For each case, the configuration of this type contributes an amount

$$\exp\left(N\alpha I(n\beta)\right) \tag{2.12}$$

in Eq. (2.5), which indicates that the contribution from RS2 is

$$\begin{aligned}
\langle Z^n\rangle &= 2^N \exp\left(nN\alpha I(n\beta)\right) \\
&= \exp N\left(\log 2 + \alpha\log\left(\cosh\frac{n\beta}{2}\right)\right).
\end{aligned} \tag{2.13}$$

- 1RSB

  In 1RSB, n replicas are assumed to be equally assigned to n/m states B, where m is an aliquot of n. Therefore, for a given B, we have

$$\sum_{\mu=1}^{n}\delta_{BA_\mu} = \left\{\begin{array}{ll} m & \text{when } B \text{ is one of the } n/m \text{ occupied states,} \\ 0 & \text{otherwise.} \end{array}\right. \tag{2.14}$$

The number of ways to select n/m out of 2^N states equally assigning n replicas to these n/m states is

$$\frac{(2^N)!}{(2^N - n/m)!} \times \frac{n!}{m^{n/m}} \sim 2^{\frac{n}{m}N}. \tag{2.15}$$

For each case, the configuration of this type contributes an amount

$$\exp\left(\frac{n}{m}N\alpha I(m\beta)\right) \tag{2.16}$$

in Eq. (2.5). Taking all the possible values of m into account, the contribution from 1RSB can be summarized as

$$\begin{aligned}
\langle Z^n\rangle &= \sum_m 2^{\frac{n}{m}N}\exp\left(\frac{n}{m}N\alpha I(m\beta)\right) \\
&= \sum_m \exp\left(\frac{n}{m}N\left[\log 2 + \alpha I(m\beta)\right]\right) \\
&\sim \exp\left(\operatorname*{extr}_m\left\{\frac{n}{m}N\left[\log 2 + \alpha I(m\beta)\right]\right\}\right).
\end{aligned} \tag{2.17}$$

In the last expression, we have replaced the summation over m with the extremization with respect to m ($\operatorname{extr}_m\{\cdots\}$), which is hopefully valid for large N, analytically continuing the expression with respect to m from the natural numbers to the real numbers. The extremization with respect to m yields the condition

$$\log 2 + \alpha I(m\beta) = \alpha m\beta I'(m\beta), \tag{2.18}$$

which implies that the moment is expressed as

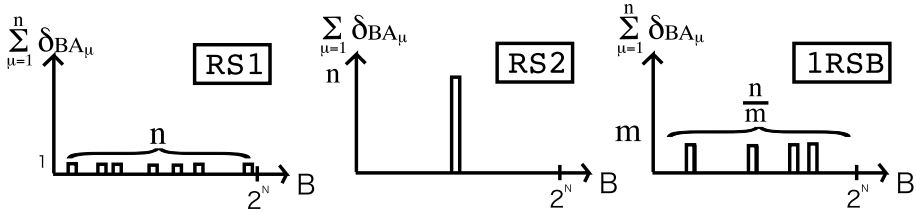$$\langle Z^n\rangle = \exp\left[nN\alpha\beta I'(m_c\beta)\right], \tag{2.19}$$

Fig. 1. The configuration of replicas for each solution. In RS1, n replicas are distributed in different states. In RS2, n replicas are concentrated in one state. In 1RSB, n replicas are equally assigned to n/m states. For 1RSB, however, this figure does not directly correspond to the solution because the critical value of m is not necessarily an integer and can be larger than n for $\beta > \beta_c$.

where m_c is the solution of Eq. (2·18). Equation (2·6) indicates that m_c can be represented as $m_c = \beta_c/\beta$, where the critical inverse temperature $\beta_c > 0$ is determined by

$$\log 2 + \alpha \left(\log \left(\cosh \frac{\beta_c}{2} \right) - \frac{\beta_c}{2} \tanh \frac{\beta_c}{2} \right) = 0. \qquad (2·20)$$

This provides a simple expression of the 1RSB solution as

$$\begin{aligned} \langle Z^n \rangle &= \exp \left(nN\alpha\beta I'(\beta_c) \right) \\ &= \exp \left(\frac{nN\alpha\beta}{2} \tanh \frac{\beta_c}{2} \right). \end{aligned} \qquad (2·21)$$

The configurations of the replicas assumed for RS1, RS2 and 1RSB are pictorially presented in Fig. 1. It should be emphasized here that the above three solutions are derived for $n = 1, 2, \cdots$, assuming N to be sufficiently large. However, it is likely that the obtained expressions hold for real n as well. Therefore, in the conventional analysis, the replica trick given by Eq. (2·4) is carried out, selecting one possibly relevant solution of the three, which is hopefully valid for large N.

The existing prescription for selecting the relevant solution is as follows.[19] For small $n > 0$, RS1, RS2 and 1RSB are ordered as RS2 > RS1 ≥ 1RSB (RS1 = 1RSB holds for $\beta = \beta_c$) with respect to their amplitudes (Fig. 2). The contribution from RS2, however, converges to 2^N rather than unity for $n \to 0$, and therefore, the replica trick leads to divergence. Hence, this solution is discarded. After excluding this solution, the leading contribution always comes from RS1. This guarantees a finite limit in Eq. (2·4).

Actually, the result obtained from this solution is correct in the case of high temperatures satisfying $0 < \beta < \beta_c$. However, this solution becomes invalid for low temperatures, i.e. $\beta > \beta_c$, for which the correct result is provided by 1RSB. It may be worth noting that the value of m_c in this low temperature case is in the interval "$[0,1]$" when $n \to 0$, which is out of the ordinary range, $1 \le m_c \le n$ for $n = 1, 2, \cdots$. This prescription for taking the $n \to 0$ limit has been empirically justified for a family of REMs because it reproduces the correct results that can be obtained using other schemes in the limit $n \to 0$.[19] However, the two issues discussed below require further investigation.

The first issue concerns the reason for expurgating RS2. Carlson's theorem, which guarantees the uniqueness of the analytic continuation from the natural numbers $n \in \mathbf{N}$ to the complex numbers $n \in \mathbf{C}$, might be useful for solving this problem.[22),23)] Unlike other systems, such as the Sherrington-Kirkpatrick model[3)] and the original REM,[19)] the modified moments of the current DREM $\left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N}$, which are extended from $n \in \mathbf{N}$ to $n \in \mathbf{C}$, satisfy the inequality

$$\left| \left\langle \left(e^{-M\beta/2} Z \right)^n \right\rangle^{1/N} \right| \leq \left\langle \left(e^{-M\beta/2} Z \right)^{\mathrm{Re}(n)} \right\rangle^{1/N}$$

$$= \left\langle \left(\sum_{A=1}^{2^N} \exp\left[-\beta(\epsilon_A + M/2) \right] \right)^{\mathrm{Re}(n)} \right\rangle^{1/N}$$

$$\leq \left\langle \left(\sum_{A=1}^{2^N} 1 \right)^{\mathrm{Re}(n)} \right\rangle^{1/N} = 2^{\mathrm{Re}(n)} < O\left(\exp\left[\pi|n| \right] \right), \quad (2\cdot22)$$

for $\mathrm{Re}(n) \geq 0$ and any finite natural numbers N, because $\epsilon_A + M/2$ is bounded from below by 0. Suppose that we could construct another extension $\psi(n; N)$ ($n \in \mathbf{C}$) that satisfies the growth condition, $\psi(n; N) < O\left(\exp\left[\pi|n| \right] \right)$,[*)] and is identical to $\left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N}$ for all the natural numbers, $n = 1, 2, \cdots$. This would imply that the similar inequality $\left| \psi(n; N) - \left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N} \right| \leq |\psi(n; N)| + \left| \left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N} \right| < O\left(\exp[\pi|n|] \right)$ holds for $\mathrm{Re}(n) \geq 0$, and that the difference $\left| \psi(n; N) - \left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N} \right|$ vanishes $\forall n \in \mathbf{N}$, as $\psi(n; N)$ and $\left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N}$ coincide $\forall n \in \mathbf{N}$. Then, Carlson's theorem (Theorem 5.81 on page 186 of Ref. 22)) ensures that $\left| \psi(n; N) - \left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N} \right|$ is identical to 0, implying that $\psi(n; N)$ and $\left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N}$ are identical. Therefore, analytic continuation of $\left\langle (e^{-M\beta/2} Z)^n \right\rangle^{1/N}$ from natural $n \in \mathbf{N}$ to complex $n \in \mathbf{C}$ can be uniquely determined. Because $e^{-M\beta/2}$ is a non-vanishing constant, this means that the analytic continuation of the moment $\langle Z^n \rangle^{1/N}$ is also unique. However, this does not guarantee that the obtained continuation remains analytic with respect to n in the limit $N \to \infty$, because certain phase transitions may occur. Actually, for high temperatures, $\beta < \beta_c$, a phase transition between RS1 and RS2 occurs in the region $n \geq 1$ in the limit $N \to \infty$. In such cases, it seems reasonable to select RS1, which is dominant for $n \gtrsim 1$, as the relevant solution for $n \to 0$, because $n = 1$ is the closest to $n = 0$ among all the natural numbers, for which correct evaluation of the moment can be carried out. This recipe successfully reproduces the correct result for the high temperature region, $0 < \beta < \beta_c$. However, this is still not fully satisfactory because RS2 becomes dominant $\forall n \in \mathbf{N}$ in the case $\beta > \beta_c$, and, therefore, it should be selected as the relevant solution for $n \to 0$, unless analyticity is lost in the range

---

[*)] This condition is necessary to exclude a trivial multiplicity caused by the addition of certain analytic functions that vanish at all the natural numbers, $n = 1, 2, \cdots$, such as $\sin(\pi n)$.
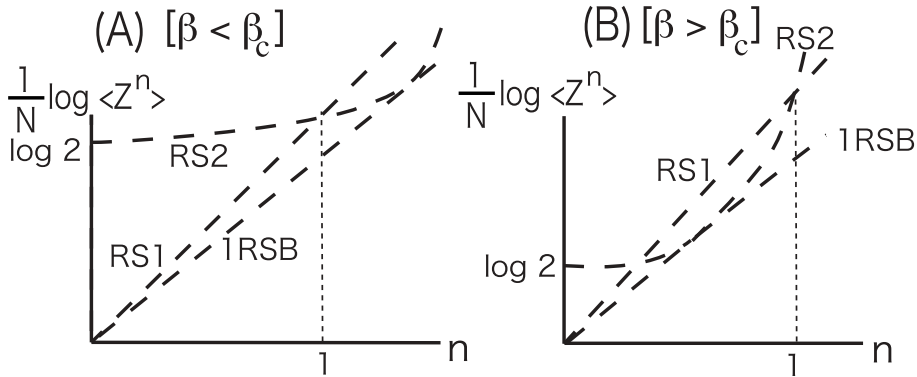
Fig. 2. Solutions obtained under the RS1, RS2 and 1RSB assumptions. RS1 and RS2 are identical at $n = 1$, while RS2 contacts 1RSB at $n = \beta_c/\beta = m_c$. (a) For $\beta < \beta_c$, the contact point is in the region of $n > 1$. (b) For $\beta > \beta_c$, contrastingly, it is located in $0 < n < 1$.

$0 < n < 1$ for $N \to \infty$, which, unfortunately, would lead to an incorrect result. Hence, a certain phase transition must occur at a critical replica number $0 < n_c < 1$ for $\beta > \beta_c$. It is worth emphasizing that determining the method for selecting the dominant solution for $0 < n < n_c$ is nontrivial in this case, because no natural numbers for which the moment can be correctly evaluated exist in the region $0 < n < 1$. However, to the knowledge of the authors, such a phase transition with respect to n has not yet been fully examined for most disordered systems.[23),24)] This might be because to this time attention has been paid mainly to the final results in the limit $n \to 0$. However, detailed analysis of phase transitions of this type may soon be needed, as the replica calculation for non-vanishing n has recently begun to be employed in problems related to IP[26),30)] and analysis of certain dynamics involved with multiple time scales.[31)]

The second issue regards the origin of 1RSB. In conventional analysis in the limit $n \to 0$,[19)-21),29)] this solution is introduced by modifying RS1 in order to keep the entropy of the correct solution non-negative for $\beta > \beta_c$. It would appear that 1RSB originates from RS1. However, at least for positive n, this association seems unlikely, because the two solutions cross only at $n = 0$ (see Fig. 2). Therefore, it is impossible to relate the origin of 1RSB to RS1 for large n, from which the solutions for smaller n are extrapolated. Contrastingly, 1RSB contacts RS2 at $n = m_c$, implying that 1RSB bifurcates from RS2. As the contact point is located in the range $0 < n < 1$ for $\beta > \beta_c$, this seems to be consistent with the extension of the aforementioned possible phase transition at $n = n_c$ [see Fig. 2 (b)]. Nevertheless, such a scenario cannot be so easily accepted, as RS2 dominates 1RSB even below the contact point; a mere contact does not change the dominance relationship between the two solutions.

It may help in resolving these problems to analyze DREM employing a completely different methodology. In the next section, we present a scheme to calculate the moments of DREM as a step towards clarifying the mysteries of RM. The proposed method is powerful enough to evaluate the moments in the right half of the complex plane, $\mathrm{Re}(n) > 0$, for arbitrary finite N, without using the replica trick.

§3. Direct calculation of the moments for the grand canonical DREM (GCDREM)

3.1. *General formula*

In order to introduce a novel scheme for calculating the moments of the partition function of the DREM, we first rewrite the partition function using the occupation numbers n_i $(i = 0, 1, 2, \cdots, M)$ as

$$Z = \sum_{A=1}^{2^N} \exp\left(-\beta\epsilon_A\right) = \sum_{i=0}^{M} n_i \exp\left(-\beta E_i\right) = \omega^{-\frac{M}{2}} \sum_{i=0}^{M} n_i \omega^i \quad (\omega \equiv e^{-\beta}). \quad (3\cdot1)$$

This means that the partition function of the DREM can be completely determined by the set of occupation numbers $\{n_i\}$, in which details of the energy configuration $\{\epsilon_A\}$ are ignored. This makes it possible to evaluate the moments $\langle Z^n \rangle$ directly from $\{n_i\}$, without referring to the full energy configuration $\{\epsilon_A\}$. This significantly reduces the necessary cost for computing the partition function when the calculation is performed numerically.

Two methods are known for generating $\{n_i\}$. The straightforward method is to count n_i independently, drawing the 2^N energy states from Eq. (2·1). The system obtained in this manner is referred to as the canonical discrete random energy model (CDREM). Although this yields a rigorously correct realization of the DREM that satisfies the constraint $\sum_{i=0}^{M} n_i = 2^N$, it requires 2^N steps to count $\{n_i\}$ and, hence, is computationally difficult. In order to resolve this difficulty in numerical experiments, Moukarzel and Parga[20], [21] proposed the grand canonical version of the discrete random energy model (GCDREM).[*] In the GCDREM, the occupation numbers are independently determined using the Poisson distribution

$$P(n_i) = e^{-\gamma_i}\frac{\gamma_i^{n_i}}{n_i!}, \quad (\gamma_i = 2^N P(E_i)) \quad (3\cdot2)$$

where γ_i is the average occupation number. The greatest advantage of the GCDREM is that it allows for the drastic reduction of the necessary computational cost for generating $\{n_i\}$ from 2^N to $M + 1$. One possible drawback of this model is that the constraint $\sum_{i=0}^{M} n_i = 2^N$ is only satisfied on average, $\left\langle \sum_{i=0}^{M} n_i \right\rangle = 2^N$. This implies that this method does not correspond strictly to the original model. However, the RM-based calculation indicates that thermodynamic properties of the GCDREM become identical to those of the CDREM as $N \to \infty$. This calculation is given in Appendix B. It can be shown that the difference rapidly vanishes as N becomes large and they are almost indistinguishable even for $N = 3$, as shown in Appendix C. In addition, this version of the DREM has another advantage in analytic calculations, because the summation can be carried out independently, as is shown below.

---

[*] Employment of the GCDREM is not essential to reduce the numerical cost. We have discovered a scheme for generating the CDREM on a time scale similar to that for the GCDREM. This scheme is given in Appendix C.

In the GCDREM, the moments are expressed as

$$\langle Z^n \rangle = \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \cdots \sum_{n_M=0}^{\infty} P(n_0)P(n_1)\cdots P(n_M)Z^n$$

$$= \omega^{-\frac{nM}{2}} \lim_{\epsilon \to 0} \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \cdots \sum_{n_M=0}^{\infty} P(n_0)P(n_1)\cdots P(n_M) \left(\sum_{i=0}^{M} n_i \omega^i + \epsilon \right)^n, (3\cdot3)$$

where an infinitesimal constant $\epsilon > 0$ is introduced in order to keep Z positive, even when all the occupation numbers vanish. This makes it possible to employ the identity for a positive number c

$$c^n = \frac{\int_H (-\rho)^{-n-1} e^{-c\rho} d\rho}{\tilde{\Gamma}(-n)} \quad [\text{with } c > 0, \tilde{\Gamma}(n) \equiv -2i \sin n\pi \Gamma(n)] \qquad (3\cdot4)$$

(whose integration contour is shown in Fig. 3) for evaluating the moment as

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\tilde{\Gamma}(-n)} \lim_{\epsilon \to 0} \int_H (-\rho)^{-n-1} e^{-(\sum_{i=0}^{M} n_i \omega^i + \epsilon)\rho} d\rho$$

$$= \frac{\omega^{-\frac{nM}{2}}}{\tilde{\Gamma}(-n)} \lim_{\epsilon \to 0} \int_H (-\rho)^{-n-1} e^{-\epsilon\rho} \left(\sum_{n_0=0}^{\infty} P(n_0)e^{-n_0\rho} \right) \left(\sum_{n_1=0}^{\infty} P(n_1)e^{-n_1\omega\rho} \right)$$

$$\cdots \left(\sum_{n_M=0}^{\infty} P(n_M)e^{-n_M\omega^M\rho} \right) d\rho. \qquad (3\cdot5)$$

It is worth noting that the summation in this expression can be carried out independently as

$$\sum_{n_i=0}^{\infty} P(n_i)e^{-n_i\omega^i\rho} = e^{-\gamma_i} \sum_{n_i=0}^{\infty} \frac{1}{n_i!}(\gamma_i e^{-\omega^i\rho})^{n_i} = \exp\left[-(1 - e^{-\omega^i\rho})\gamma_i\right]. \qquad (3\cdot6)$$

Therefore, the moment can be summarized as

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\tilde{\Gamma}(-n)} \lim_{\epsilon \to 0} \int_H (-\rho)^{-n-1} \exp\left[-\epsilon\rho - \sum_{i=0}^{M}(1 - e^{-\omega^i\rho})\gamma_i \right] d\rho. \qquad (3\cdot7)$$

Because this is convergent for $\text{Re}(n) > 0$, the following expression gives the analytic continuation of the moment to the right half complex plane of n:

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\tilde{\Gamma}(-n)} \int_H (-\rho)^{-n-1} \exp\left[-\sum_{i=0}^{M}(1 - e^{-\omega^i\rho})\gamma_i \right] d\rho. \qquad (3\cdot8)$$

3.2.   *Thermodynamic limit*

Using Eq. (3·8), one can analytically examine the behavior of the moments in the thermodynamic limit $N, M \to \infty$, keeping $\alpha = M/N$ finite. For this purpose,
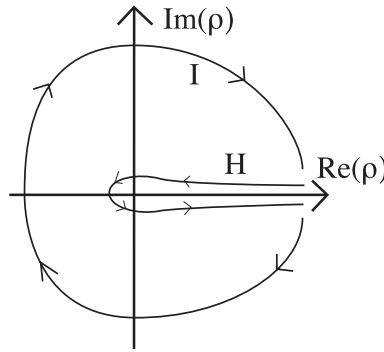
Fig. 3. The integration contours.

we first convert the contour integration in the expression into an integration on the real axis for $p - 1 < \text{Re}(n) < p$, where p is an arbitrary natural number. Further, we define the function

$$f(\rho) \equiv \exp\left[-\sum_{i=0}^{M}(1 - e^{-\omega^i \rho})\gamma_i\right] \equiv \sum_{j=0}^{\infty} f_i \rho^i \tag{3.9}$$

and a series of truncated summations as

$$f^{(p)}(\rho) \equiv \sum_{j=0}^{p-1} f_i \rho^i, \tag{3.10}$$

which satisfy the identity

$$\int_H (-\rho)^{-n-1} f^{(p)}(\rho) d\rho = \int_{H+I} (-\rho)^{-n-1} f^{(p)}(\rho) d\rho = 0 \tag{3.11}$$

for $p - 1 < \text{Re}(n)$, because the contribution from the contour I vanishes. Using this identity, the moment can be rewritten as

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\tilde{\Gamma}(-n)} \int_H (-\rho)^{-n-1}[f(\rho) - f^{(p)}(\rho)]d\rho. \tag{3.12}$$

Equation (3.11) guarantees that the infrared divergence is removed for $\text{Re}(n) < p$ in this expression. Therefore, the moment can be rewritten for $p - 1 < \text{Re}(n) < p$ as

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\Gamma(-n)} \int_0^\infty \rho^{-n-1}[f(\rho) - f^{(p)}(\rho)]d\rho, \tag{3.13}$$

where the function $\tilde{\Gamma}$ is replaced by the ordinary gamma function Γ.

As we have particular interest in the case $p = 1$ (i.e. $0 < \text{Re}(n) < 1$), let us focus on the behavior of the expression

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\Gamma(-n)} \int_0^\infty \rho^{-n-1}[f(\rho) - 1]d\rho$$

$$= \frac{\omega^{-\frac{nM}{2}}}{\Gamma(-n)} \int_0^\infty \rho^{-n-1} \Big[\exp\Big\{ -\sum_{i=0}^{M}(1 - e^{-\omega^i \rho})\gamma_i \Big\} - 1 \Big] d\rho. \qquad (3\cdot14)$$

To examine the behavior of the thermodynamic limit, it is convenient to introduce the new variables

$$x \equiv \frac{i}{N\alpha}, \quad y \equiv \frac{1}{N\alpha\beta} \ln \rho \qquad (3\cdot15)$$

in place of i and ρ. This yields the expression

$$\langle Z^n \rangle = \frac{\omega^{-\frac{nM}{2}}}{\Gamma(-n)} \int_{-\infty}^\infty e^{N\mathcal{G}(y)} dy \qquad (3\cdot16)$$

for the moments, where

$$\begin{cases} \mathcal{G}(y) & = -n\alpha\beta y + \dfrac{1}{N}\log(1 - e^{-\mathcal{F}(y)}), \\[2mm] \mathcal{F}(y) & = \displaystyle\int_0^1 e^{N\mathcal{H}(x,y)} dx, \\[2mm] \mathcal{H}(x,y) & = (1-\alpha)\log 2 + \alpha H(x) + \dfrac{1}{N}\log(1 - e^{-e^{N\alpha\beta(y-x)}}), \end{cases} \qquad (3\cdot17)$$

and

$$H(x) = -x\log x - (1-x)\log(1-x). \qquad (3\cdot18)$$

Here, we have replaced the summation with an integration. This is valid when both M and N are sufficiently large.

For further analysis, the identity

$$g(u) \equiv \frac{1}{N}\log\left(1 - e^{-e^{Nu}}\right) \to \begin{cases} 0 & (u \geq 0) \\ u & (u < 0) \end{cases}$$
$$= u\theta(-u), \qquad (3\cdot19)$$

which holds for large N, may be useful. The shape of this function is displayed in Fig. 4. Because this becomes singular at $u = 0$, a phase transition may occur in n space as $N \to \infty$. We examine this possibility below.

Equation (3·19) imply that the function $\mathcal{H}$ can be expressed as

$$\mathcal{H}(x,y) = (1-\alpha)\log 2 + \alpha H(x) + \alpha\beta(y-x)\theta(y-x) \qquad (3\cdot20)$$

in the thermodynamic limit. As a function of x, this exhibits three types of behavior, depending on the value of y, as shown in Fig. 5. Employing the saddle point method,
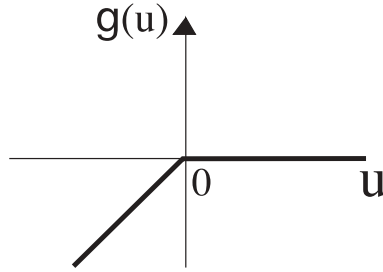
$$g(u)$$

Fig. 4. The shape of $g(u)$ for $N \to \infty$. This is not analytic at $u = 0$.

the maximum of $\mathcal{H}(x, y)$ given y yields $\mathcal{F}(y)$ for large N. This is

$$\lim_{N \to \infty} \frac{1}{N} \log \mathcal{F}(y) = \begin{cases} \frac{1}{2} & \left(y > \frac{1}{2}\right), \\ (1 - \alpha) \log 2 + \alpha H(x) & \left(x_c < y < \frac{1}{2}\right), \\ (1 - \alpha) \log 2 + \alpha \tilde{H}(\beta) + \alpha \beta y & (y < x_c), \end{cases} \tag{3.21}$$

where x_c is a solution of

$$H'(x_c) = \beta, \tag{3.22}$$

whose behavior is shown in Fig. 6. Here, the function $\tilde{H}$ is the Legendre transform of the function H:

$$\tilde{H}(\beta) = \log\left(2 \cosh \frac{\beta}{2}\right) - \frac{\beta}{2}. \tag{3.23}$$

The function $\mathcal{G}$ becomes

$$\mathcal{G}(y) = -n\alpha\beta y + \frac{1}{N} \log \mathcal{F}(y)\theta(-\mathcal{F}(y)) \tag{3.24}$$

for large N, which directly controls the behavior of the moment in Eq. (3.16). This behavior depends strongly on the relation between x_c and $x^* = \left(1 - \tanh \frac{\beta_c}{2}\right)/2$, which satisfies the condition

$$(1 - \alpha) \log 2 + \alpha H(x^*) = 0. \tag{3.25}$$

- (A) $x^* < x_c$
  As x_c and x^* are defined in Eqs. (3.22) and (3.25), respectively, the condition $x^* < x_c$ can be written

$$(1 - \alpha) \log 2 + \alpha H(x_c) > 0. \tag{3.26}$$

This is satisfied for $\beta < \beta_c$, i.e. in the high temperature phase. Notice that for $\alpha \leq 1$, as Eq. (2.20) does not have a positive solution, this is always satisfied independently of β. Then, Eq. (3.24) can be written

$$\mathcal{G}(y) = -n\alpha\beta y + [(1 - \alpha) \log 2 + \alpha \tilde{H}(\beta) + \alpha\beta y]\theta(y_c - y), \tag{3.27}$$
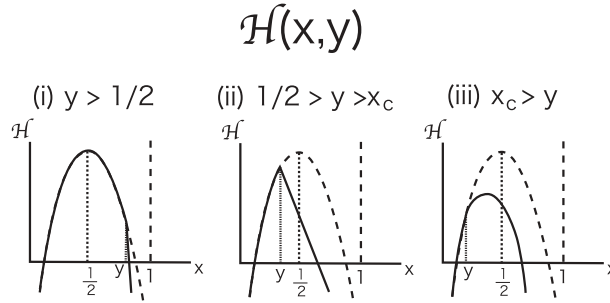
$\mathcal{H}(x,y)$

Fig. 5.   The schematic shape of $\mathcal{H}(x, y)$ for $N \to \infty$. It depends on the value of y.
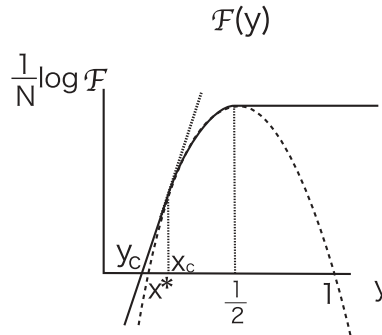
$\mathcal{F}(y)$

Fig. 6.   The schematic shape of $\mathcal{F}(y)$ for $N \to \infty$ obtained from the maximal value of $\mathcal{H}(x, y)$.

which is shown in Fig. 7. Here, y_c is defined by $\mathcal{F}(y_c) = 0$ (Fig. 6), which yields

$$y_c = -\frac{1}{\alpha\beta}[(1 - \alpha)\log 2 + \alpha\tilde{H}(\beta)]. \tag{3\cdot28}$$

The moments are then calculated as

$$\langle Z^n \rangle = -N\alpha\beta\frac{\omega^{-\frac{nM}{2}}}{\Gamma(-n)} \left[\int_{-\infty}^{y_c} \exp\left(1 - \alpha\right)\log 2 + \alpha\tilde{H}(\beta) + \alpha\beta(1 - n)y dy \right.$$
$$\left. + \int_{y_c}^{\infty} \exp\left(-n\alpha\beta y\right)dy \right]$$
$$\sim e^{\alpha\beta nN(\frac{1}{2} - y_c)} \left[\frac{1}{\Gamma(1 - n)} + \frac{n}{\Gamma(2 - n)} \right]. \tag{3\cdot29}$$

Therefore, the asymptotic behavior of the moments are given by

$$\lim_{N \to \infty} \frac{1}{N}\log\langle Z^n \rangle = \alpha\beta nN\left(\frac{1}{2} - y_c\right)$$
$$= n\left(\log 2 + \alpha\log\cosh\frac{\beta}{2}\right) \tag{3\cdot30}$$

in this phase. This is consistent with RS1, as can be seen from Eq. (2·10).

- (B) $x^* > x_c$
  The condition $x^* > x_c$ is satisfied for $\beta > \beta_c$, which may correspond to the low
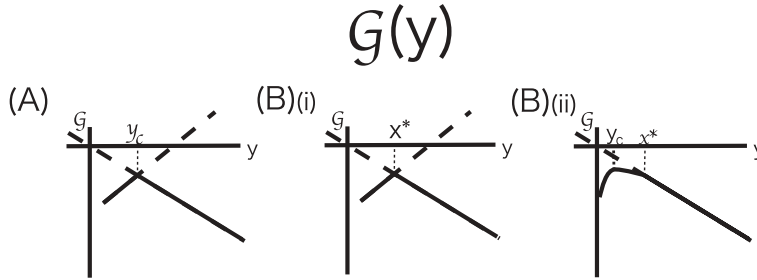
$$\mathcal{G}(y)$$

Fig. 7. Schematic figure of $\mathcal{G}(y)$. The maximal value determines the behavior of $\langle Z^n \rangle$.

temperature phase. The function $\mathcal{G}(y)$ has the form

$$\mathcal{G}(y) = -n\alpha\beta y + [(1-\alpha)\log 2 + \alpha H(y)]\theta(x^* - y). \tag{3.31}$$

Its behavior of which is further classified into two cases, depending on the replica number n.

– (i) $H'(x^*) > n\beta$

Because x^* is located in $x_c < x^* < \frac{1}{2}$, there exists a critical replica number $0 < n_c < 1$, which is characterized by

$$n_c \equiv \frac{1}{\beta} H'(x^*) = \frac{\beta_c}{\beta} = m_c. \tag{3.32}$$

The condition $H'(x^*) > n\beta$ is satisfied for $n < n_c$. The function $\mathcal{G}(y)$, shown in Fig. 7, is maximized at $y = x^*$. Therefore, the moments can be expressed as

$$\lim_{N \to \infty} \frac{1}{N} \log \langle Z^n \rangle = n\alpha\beta \left(\frac{1}{2} - x^* \right)$$
$$= \frac{n\alpha\beta}{2} \tanh \frac{\beta_c}{2}. \tag{3.33}$$

This behavior is identical to that of the 1RSB predicted by the RM, as can be seen in Eq. (2.21).

– (ii) $H'(x^*) < n\beta$

The condition $H'(x^*) < n\beta$ is satisfied for $n > n_c$. The function $\mathcal{G}(y)$ is maximized not at $y = x^*$ but at $y = y_c$. Therefore, the moments can be asymptotically expressed as

$$\lim_{N \to \infty} \frac{1}{N} \log \langle Z^n \rangle = \log 2 + \alpha \log \cosh \frac{n\beta}{2}. \tag{3.34}$$

This coincides with the behavior of RS2 obtained using RM, as can be seen in Eq. (2.13).

The results are summarized in Fig. 8. In the high temperature phase, $\beta < \beta_c$, the behavior of the moments is simple, being expressed by RS1 of RM. In the low temperature phase, $\beta > \beta_c$, the behavior of the moments has two possibilities,
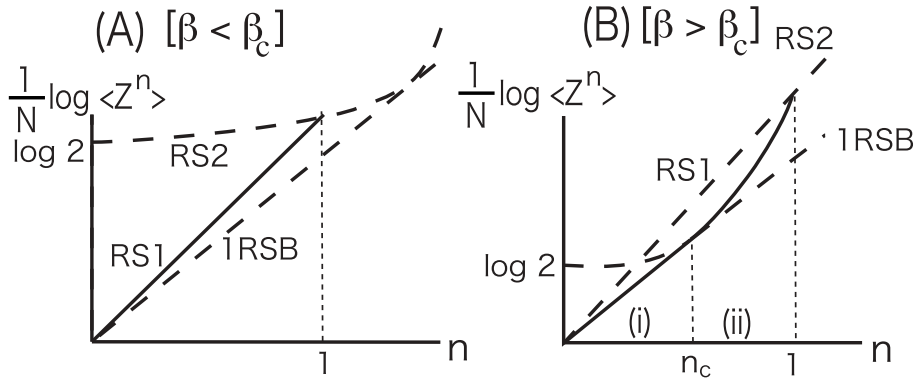
Fig. 8. The behavior of the moments in the thermodynamic limit obtained from the exact expression (solid curves). This line is shown only in the region $0 < n < 1$ because the expression Eq. (3·14) is valid for $0 < \mathrm{Re}(n) < 1$, while the original expression Eq. (3·8) is valid for $0 < \mathrm{Re}(n)$. In the high temperature phase, $\beta < \beta_c$, the behavior of the moments is simple and corresponds to the result obtained from RS1. In the low temperature phase, $\beta > \beta_c$, the behavior of the moments has interesting properties. The moments correspond to those obtained from RS2 for $n > n_c \equiv \beta/\beta_c = m_c$ and correspond to those obtained from 1RSB for $n < n_c$. A phase transition occurs at $n = n_c$.

depending on n. More specifically, in the limit $N \to \infty$, the moments approach RS2 for $n > n_c$, whereas 1RSB represents the correct behavior for $n < n_c$. This means that there exists a phase transition in the space of the replica number at $n = n_c$. In conclusion, these are consistent with the known results obtained using RM in the limit $n \to 0$.[19)–21)]

3.3. *Numerical validation*

Equation (3·8) is formulated as a two-dimensional summation with respect to $\rho \in \boldsymbol{C}$ and $i = 0, 1, \cdots, M$ and is numerically tractable. This means that Eq. (3·8) or Eq. (3·16) can be utilized to numerically examine the behavior of GCDREM for a finite system size N, and to study how fast the results obtained for $N \to \infty$ become relevant as N increases.

Figure 9 plots the logarithm of $\langle Z^n \rangle$ calculated using Eq. (3·16) and $\langle Z^n \rangle$ evaluated numerically from 10, 1000 and 100,000 experiments, with $N = 10$ for GCDREM. One can see that the data from the numerical experiments converge to the results obtained from Eq. (3·16). This verifies that our expression provides accurate values for the moments, even for a finite system size.

We next compare the results from our expression and the RM in Fig. 10. At high temperatures, our result is consistent with RS1 as expected. The difference is negligible for all of the range $0 < n < 1$ even at $N = 10$. At low temperatures, our result fits RS2 for larger $n > n_c$, while 1RSB exhibits excellent consistency for smaller $n < n_c$. There is a slight difference between our expression and 1RSB for $N = 10$. The difference, however, becomes indistinguishable for $N = 100$. This strongly indicates that there occurs a phase transition between RS2 and 1RSB at $n = n_c$ in the limit $N \to \infty$.
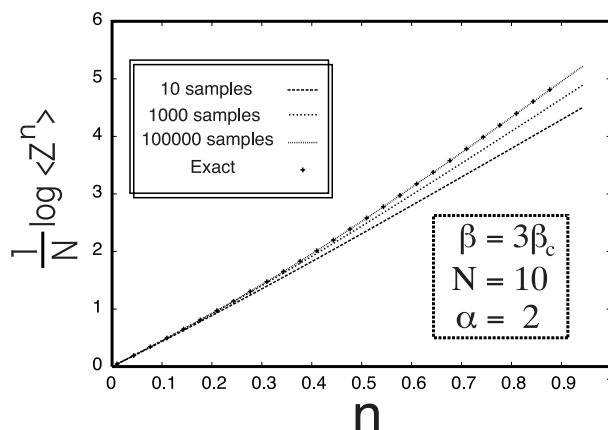
Fig. 9. The logarithm of $\langle Z^n \rangle$ calculated from Eq. (3·16) and data obtained from numerical experiments of Eq. (3·1) using 10, 1000 and 100,000 samples for GCDREM. The system size is $N = 10$. The points indicated as "Exact" are the results obtained from Eq. (3·16).

§4.   Origin of 1RSB: Extreme value statistics

The results found in the preceding two sections indicate that the DREM exhibits a phase transition with respect to the replica number n at a certain critical point $n_c \in [0, 1]$ when the temperature is sufficiently low. In this section, we discuss how this transition can be understood in the framework of the RM. A formalism analogous to that previously introduced for examining the domain size distribution of multi-layer perceptrons is useful for this purpose.[30]

Because the partition function of the DREM typically scales exponentially with respect to N, we first express this dependence as

$$Z \sim \exp\left[-N\alpha\beta\left(y - \frac{1}{2}\right)\right], \tag{4·1}$$

where $y - \frac{1}{2}$ represents the free energy normalized by the scale of the energy amplitude M for a given realization $\{\epsilon_A\}$ [$A = 1, 2, \cdots, 2^N$) (or $\{n_i\}$ ($i = 0, 1, \cdots, M$)]. Clearly, y is a random variable. Let us assume that the probability distribution of y, $\mathcal{P}(y)$, scales as

$$\mathcal{P}(y) \sim \exp\left[-Nc(y)\right], \tag{4·2}$$

where $c(y) \sim O(1)$ for large N. Note that the inequality

$$c(y) \geq 0 \tag{4·3}$$

must hold in order for $\mathcal{P}(y)$ to satisfy the normalization condition $\int dy \mathcal{P}(y) = 1$ for $N \to \infty$. Equation (4·2) indicates that the moment of the partition function can be calculated as

$$\langle Z^n \rangle \equiv \int dy \exp\left[-Nn\alpha\beta\left(y - \frac{1}{2}\right)\right] \mathcal{P}(y)$$
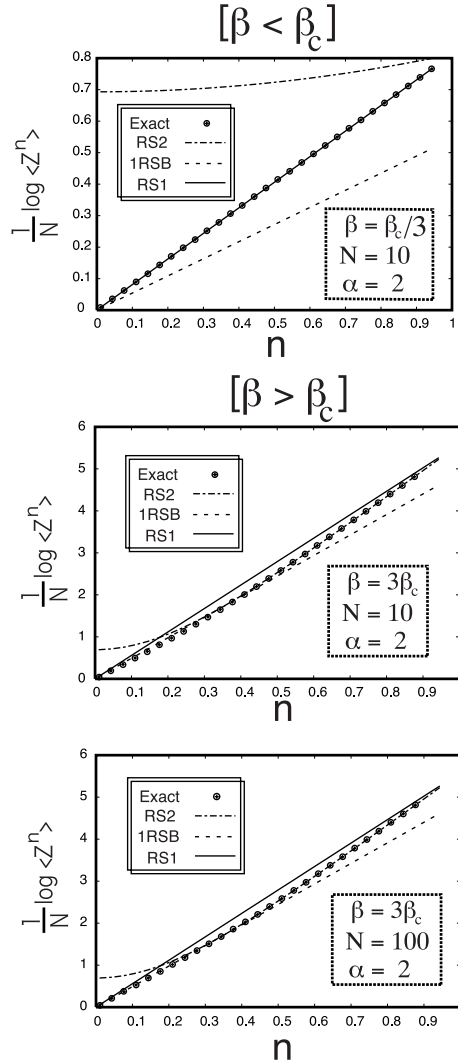
Fig. 10. The logarithm of $\langle Z^n \rangle$ obtained from our expression and the RM. The points indicated as "Exact" are the results obtained from Eq. (3·16). The lines indicated "RS1", "RS2", and "1RSB" are the results obtained from the RM, using Eqs. (2·10), (2·13) and (2·21), respectively. At high temperatures, $\beta = \beta_c/3 < \beta_c$, our result is consistent with RS1. The difference is very small even for $N = 10$. At low temperatures, $\beta = 3\beta_c > \beta$, our result fits RS2 for larger $n > n_c = \beta_c/\beta = 1/3$, while 1RSB provides good consistency for smaller $n < n_c$. There is a slight difference between the results obtained from our expression and those obtained from 1RSB at $N = 10$. The difference, however, is imperceptible for $N = 100$.

$$\sim \exp\left[N \operatorname*{extr}_{y} \left\{ -n\alpha\beta \left(y - \frac{1}{2} \right) - c(y) \right\} \right]. \qquad (4\cdot4)$$

This formula, however, may not be useful for computing $\langle Z^n \rangle$, as directly assessing $c(y)$ is rather difficult. Instead, it can be utilized to evaluate $c(y)$ from $\langle Z^n \rangle$, as

Eq. (4·4) implies that $c(y)$ can be obtained from

$$c(y) = \operatorname*{extr}_{n} \left\{ -n\alpha\beta \left(y - \frac{1}{2} \right) - \frac{1}{N} \ln \langle Z^n \rangle \right\}, \tag{4·5}$$

where $(1/N) \ln \langle Z^n \rangle$ can be computed using the RM for any real number n. This, in conjunction with the normalization constraint in Eq. (4·3), offers useful information to identify the origin of the phase transition with the aid of the RM.

For $\beta > \beta_c$, RS2 provides the dominant solution for the natural numbers $n = 1, 2, \cdots$ in the thermodynamic limit. Therefore, let us first insert this solution, from Eq. (2·13), into Eq. (4·5). This yields

$$
\begin{aligned}
c(y) &= \operatorname*{extr}_{n} \left\{ -n\alpha\beta \left(y - \frac{1}{2} \right) - \log 2 - \alpha \log \left(\cosh \frac{n\beta}{2} \right) \right\} \\
&= (\alpha - 1) \log 2 - \alpha H(y).
\end{aligned} \tag{4·6}
$$

Note that this does not satisfy the necessary condition given in Eq. (4·3) for $y > x^*$, and the critical value $y = x^*$ corresponds to $n = n_c = H'(x^*)/\beta = \beta_c/\beta = m_c$, which signals the occurrence of a phase transition at $n = n_c$ within the framework of the RM.

The following considerations concerning the energy configuration provides a plausible scenario for this transition. As has already been pointed out,[29] the partition function Z in the low temperature region, $\beta > \beta_c$, can be considered to be dominated by only the minimum energy, $\epsilon_{\min}$, of a given energy configuration $\{\epsilon_A\}$, i.e., $Z \sim \exp[-\beta\epsilon_{\min}]$. This implies that $\mathcal{P}(y)$ of the normalized free energy $y - \frac{1}{2} = -(\log Z)/(N\alpha\beta)$ is given by the distribution of $\epsilon_{\min}/M$ under the assumption that each energy level ϵ_A $(A = 1, 2, \cdots, 2^N)$ is independently drawn from an identical distribution, given in Eq. (2·1), which has already been studied in the context of *extreme value statistics* (EVS).[32] In the present case, the average occupation number of an energy level $E_i = i - \frac{1}{2}M = M\left(x - \frac{1}{2} \right)$, that is,
$\gamma_i = 2^{N-M} \left(\begin{array}{c} M \\ M/2 + E_i \end{array} \right) = \exp N\left((1 - \alpha)\log 2 + \alpha H(x) \right)$ $(i = 0, 1, \cdots, M$ and $x = i/M \in [0, 1])$, grows exponentially with respect to N for $x > x^*$. This indicates that $\mathcal{P}(y)$ is very small for $y > x^*$, because $x > x^*$ does not provide $\epsilon_{\min}/M$ for a given energy configuration $\{\epsilon_A\}$, except in very rare cases, as energy levels lower than $E_i = M\left(x - \frac{1}{2} \right)$ are included in the configuration with a very high probability. On the other hand, γ_i decreases exponentially for $0 < x < x^*$, which means that the probability of having an energy level labelled by some x in this interval in the energy configuration is low. Therefore, $\mathcal{P}(y)$ also becomes small for $0 < y < x^*$. However, the functional form of $\mathcal{P}(y)$ is not symmetric between $y > x^*$ and $0 < y < x^*$. Actually, detailed analysis of EVS[29],[32] shows that $\mathcal{P}(y)$ exhibits the asymmetric scaling forms

$$\mathcal{P}(y) \sim \begin{cases} \exp\left[-\exp\left[-N\left((\alpha - 1)\log 2 - \alpha H(y) \right) \right] \right], & (x^* < y) \\ \exp\left[-N\left((\alpha - 1)\log 2 - \alpha H(y) \right) \right], & (0 < y < x^*) \end{cases} \tag{4·7}$$

for large N, which yield a singularity at $y = x^*$ in the limit $N \to \infty$. It may be worth noting that $1 - \exp[-\mathcal{F}(y)]$, which appeared in Eq. (3·16) after inserting

Eq. (3·17) in the analysis presented in §3, corresponds to the cumulative distribution $\int_{-\infty}^{y} d\tilde{y} \mathcal{P}(\tilde{y})$. Equation (4·7) implies that the extremum point in Eq. (4·4) for large $n \ (> n_c)$ is in $0 < y < x^*$. This leads to RS2 and consistently provides the exponent of Eq. (4·6). On the other hand, the constant $y = x^*$ gives the extremum for all small $n \ (< n_c)$, which corresponds to 1RSB in the framework of RM. The intuitive implication of this is that $\langle Z^n \rangle$ is dominated by an atypically low minimum energy generated with a small probability for $n > n_c$, whereas the typical minimum energy $\epsilon_{\min} = M \left(x^* - \frac{1}{2} \right)$ provides the major contribution to $\langle Z^n \rangle$ for $n < n_c$. Thus, the origin of the phase transition can be identified as the singularity of the free energy distribution $\mathcal{P}(y)$ exhibited in the case of low temperatures, $\beta > \beta_c$.

   Our analysis also demonstrates that Carlson's theorem is not necessarily useful for validating the RM, because an analytic continuation formulated for finite N can exhibit a singularity in the limit $N \to \infty$, and therefore, taking the limit $N \to \infty$ prior to $n \to 0$ in determining the continuation on the basis of the expressions for $n = 1, 2, \cdots$, which is usually done in the RM, sometimes yields an incorrect result for $n < 1$. However, in the present system, this drawback can be overcome by taking a constraint on the free energy distribution [i.e. Eq. (4·3)] into account. This leads to the conventional 1RSB. Although the importance of the concept of EVS in the RM has been already addressed with regard to the limit $n \to 0$,[29] to our knowledge, the present analysis is the first that directly clarifies how the properties of EVS relate to the corruption of the analytic continuation in the RM, expressed as a phase transition with respect to the replica number n.

§5. Summary

   In summary, we have offered an exact expression, Eq. (3·8), of the moments of the partition function $\langle Z^n \rangle$ for the GCDREM, without employing the replica trick. This expression is valid for an arbitrary system size N and complex number n [with $\text{Re}(n) > 0$]. Simplifying the expression for the case $0 < n < 1$, we have shown that a phase transition with respect to the number of replicas n occurs at a certain critical number $n_c \in [0, 1]$ in the thermodynamic limit $N \to \infty$ when the temperature β^{-1} is sufficiently low, if the ratio $\alpha = M/N$ is greater than unity. This implies that Carlson's theorem, which guarantees the uniqueness of the analytic continuation from the expressions for $n = 1, 2, \cdots$ to those for $n \in \boldsymbol{C}$, is not necessarily useful for validating the replica method, because taking the thermodynamic limit $N \to \infty$ prior to the continuation results in an incorrect result when the analyticity is lost for $n < 1$ in the case of infinite N. However, it has also been shown that this drawback can be overcome by taking the statistical properties of the minimum energy level into account. This clarifies how 1RSB originates from a replica symmetric solution (which has been conventionally discarded) at the critical replica number n_c. We hope that the results obtained here can provide a useful insight into the remaining mysteries of the RM.

   Because Eq. (3·8) is valid throughout the right half complex plane $\text{Re}(n) > 0$ for any finite N, one can directly observe how the singularities of $\langle Z^n \rangle$ approach the real axis of n as $N \to \infty$.[23] This is sometimes referred to as Lee-Yang's scenario of

the phase transition.[33] Analysis of this point is currently under way.[34]

Acknowledgements

Appendix A
—— *Relation between the DREM and Error-Correcting Codes* ——

As was shown in Ref. 9), certain types of spin glass models can be related to error-correcting codes. Here, we show that a known evaluation scheme of the error correcting ability of a random code ensemble can be reduced to a calculation of the general moment of the partition function with respect to the DREM.[35],[36] A remark on this relationship is made in Ref. 37).

In the general scenario of error correcting codes, an N-dimensional binary vector $\boldsymbol{s} = (s_1, s_2, \cdots, s_N)$ $(s_i = 0, 1, i = 1, 2, \cdots, N)$ is encoded as a codeword $\boldsymbol{t} = (t_1, t_2, \cdots, t_M)$ $(t_i = 0, 1, i = 1, 2, \cdots, M)$ of an M $(> N)$-dimensional binary vector and transmitted via a noisy channel. Here, we concentrate on a binary symmetric channel (BSC) in which each component of the codeword t_i is independently flipped to the opposite letter in the alphabet (0 or 1) with a probability $0 < p < 1/2$. This implies that the vector $\boldsymbol{r} = \boldsymbol{t} + \boldsymbol{n}$ (mod 2) is received at the other terminal of the channel, where $\boldsymbol{n}$ is the noise vector, whose components become 1 independently with probability p and are 0 otherwise. However, as the codeword is represented somewhat redundantly, decoding $\boldsymbol{r}$ provides the correct message $\boldsymbol{s}$ with a high probability for sufficiently small p.

It is known that the performance for the task of correctly retrieving $\boldsymbol{s}$ from $\boldsymbol{r}$ becomes good when a code $\mathcal{C}$ (i.e., an invertible map $\mathcal{C} : \boldsymbol{s} \leftrightarrow \boldsymbol{t}$) is randomly constructed,[25],[35] providing a code ensemble. Let us evaluate the average probability of failing to correctly retrieve $\boldsymbol{s}$ in order to characterize the potential error correcting ability of the ensemble. We focus on the maximum likelihood (ML) decoding, which selects the codeword closest to the received vector $\boldsymbol{r}$ and returns a message vector corresponding to the codeword as an estimate $\hat{\boldsymbol{s}}$ of $\boldsymbol{s}$. ML decoding minimizes the decoding error probability $P_E(\mathcal{C})$ when codeword vectors $\boldsymbol{t}$ are equally generated at the transmission terminal.[38]

Note that the message $\boldsymbol{s}$ can be correctly identified if its codeword $\boldsymbol{t}$ is provided, because the code is constructed as an invertible map. Therefore, to evaluate $P_E(\mathcal{C})$, it is convenient to introduce an indicator function $\Delta_{\mathrm{ML}}(\boldsymbol{t}, \boldsymbol{r}|\mathcal{C})$, which returns 1 if $\boldsymbol{r}$ is not correctly decoded as the correct codeword $\boldsymbol{t}$ and 0 otherwise. Thus, the decoding error probability can be computed as

$$P_E(\mathcal{C}) = \sum_{\boldsymbol{t}, \boldsymbol{r}} P(\boldsymbol{t}|\mathcal{C}) P(\boldsymbol{r}|\boldsymbol{t}) \Delta_{\mathrm{ML}}(\boldsymbol{t}, \boldsymbol{r}|\mathcal{C}). \tag{A.1}$$

Hence, its average over the code ensemble can be represented as

$$\langle P_E(\mathcal{C})\rangle_{\mathcal{C}} = \sum_{\mathcal{C}} P(\mathcal{C}) \sum_{\boldsymbol{t}\in\mathcal{C},\boldsymbol{r}} P(\boldsymbol{t}|\mathcal{C})P(\boldsymbol{r}|\boldsymbol{t})\Delta_{\text{ML}}(\boldsymbol{t},\boldsymbol{r}|\mathcal{C}), \tag{A·2}$$

where $P(\boldsymbol{t}|\mathcal{C})$ is the probability that the codeword vector $\boldsymbol{t}$ is generated given $\mathcal{C}$, and $P(\boldsymbol{r}|\boldsymbol{t})$ is the conditional probability that $\boldsymbol{r}$ is received when $\boldsymbol{t}$ is transmitted. $P(\mathcal{C})$ is the probability that a code $\mathcal{C}$ is generated. The symbol $\sum_{\boldsymbol{t}\in\mathcal{C}}$ represents summation over the 2^N codeword vectors, given $\mathcal{C}$.

We assume that the code $\mathcal{C}$ is designed by the source coding technique such that $\boldsymbol{t}$ is uniformly generated as $P(\boldsymbol{t}|\mathcal{C}) = 1/2^N$.[35] In addition, for BSC, the conditional probability can be represented as

$$P(\boldsymbol{r}|\boldsymbol{t}) = \frac{\exp\left[-F\sum_{i=1}^{M}\left(\frac{1}{2}-\delta_{r_i,t_i}\right)\right]}{\left(2\cosh\frac{F}{2}\right)^M}, \tag{A·3}$$

where $\delta_{x,y}$ is 1 if $x = y$ and 0 otherwise, and $F = \log[(1-p)/p]$.

Unfortunately, expressing $\Delta_{\text{ML}}(\boldsymbol{t},\boldsymbol{r}|\mathcal{C})$ in a rigorously treatable form is difficult. However, Gallager's inequality,

$$\Delta_{\text{ML}}(\boldsymbol{t},\boldsymbol{r}|\mathcal{C}) \leq \left(\sum_{\boldsymbol{t}'\in\mathcal{C}\backslash\boldsymbol{t}}\left(\frac{P(\boldsymbol{r}|\boldsymbol{t}')}{P(\boldsymbol{r}|\boldsymbol{t})}\right)^{\frac{1}{1+n}}\right)^n, \tag{A·4}$$

which holds for arbitrary $n \geq 0$, offers a good upper bound.[35],[36] Here, $\sum_{\boldsymbol{t}'\in\mathcal{C}\backslash\boldsymbol{t}}$ represents a summation over the $2^N - 1$ codeword vectors $\boldsymbol{t}'$ of $\mathcal{C}$, excluding the possibility of the correct codeword $\boldsymbol{t}$. Inserting this into Eq. (A·2) provides

$$\begin{aligned}
\langle P_E(\mathcal{C})\rangle_{\mathcal{C}} &\leq \sum_{\mathcal{C}} P(\mathcal{C}) \sum_{\boldsymbol{t},\boldsymbol{r}} P(\boldsymbol{t}|\mathcal{C})P^{\frac{1}{1+n}}(\boldsymbol{r}|\boldsymbol{t})\left(\sum_{\boldsymbol{t}'\in\mathcal{C}\backslash\boldsymbol{t}} P^{\frac{1}{1+n}}(\boldsymbol{r}|\boldsymbol{t})\right)^n \\
&= \sum_{\mathcal{C}} P(\mathcal{C}) \sum_{\boldsymbol{r}} P^{\frac{1}{1+n}}(\boldsymbol{r}|\boldsymbol{0})\left(\sum_{\boldsymbol{t}'\in\mathcal{C}\backslash\boldsymbol{0}} P^{\frac{1}{1+n}}(\boldsymbol{r}|\boldsymbol{t}')\right)^n, \\
&= \sum_{r_i=\pm1} \frac{e^{\frac{F}{1+n}\sum_{i=1}^{M}\left(\frac{1}{2}-r_i\right)}}{\left(2\cosh\frac{F}{2}\right)^M} \sum_{\mathcal{C}} P(\mathcal{C})\left(\sum_{\boldsymbol{t}'\in\mathcal{C}\backslash\boldsymbol{0}} e^{-\frac{F}{1+n}\sum_{i=1}^{M}\left(\frac{1}{2}-\delta_{r_i,t_i'}\right)}\right)^n, \tag{A·5}
\end{aligned}$$

where we have carried out the gauge transformation $\boldsymbol{t}\to\boldsymbol{0}$, $\boldsymbol{t}'-\boldsymbol{t}\to\boldsymbol{t}'$ and $\boldsymbol{r}-\boldsymbol{t}\to\boldsymbol{r}$, assuming that any code $\mathcal{C}$ contains the zero codeword $\boldsymbol{t}=\boldsymbol{0}$.[35] Minimizing the final expression with respect to $n \geq 0$, we can obtain the tightest bound on the average decoding error probability.

Here, we discuss the fact that $E(\boldsymbol{t}) = \sum_{i=1}^{M}\left(\frac{1}{2}-\delta_{r_i,t_i}\right)$ in Eq. (A·5) satisfies Eq. (2·1) independently of $\boldsymbol{r}$ when each codeword $\boldsymbol{t}$ is generated with an equal probability in the ensemble in such a manner that each component is independently selected from an identical unbiased distribution $P(t_i = 1) = P(t_i = 0) = 1/2$. This

condition holds for the random code ensemble,[25),35)] although non-trivial correlations of the 'energy' $E(\boldsymbol{t})$ between 'states' $\boldsymbol{t}$ arise. This causes the energy distribution to differ from Eq. (2·1) in the case of practical linear codes. Therefore, for the current random code ensemble, Eq. (A·5) can be simplified as

$$\langle P_E(\mathcal{C})\rangle_{\mathcal{C}} \leq \left(\frac{\cosh\frac{F}{2(1+n)}}{\cosh\frac{F}{2}}\right)^M \left\langle\left(\sum_{A=1}^{2^N-1}\exp\left[-\frac{F}{1+n}\epsilon_A\right]\right)^n\right\rangle, \qquad (\text{A·6})$$

where $\langle\cdots\rangle$ denotes the configurational average with respect to the 'energy states' ϵ_A ($A = 1, 2, \cdots, 2^N - 1$), following Eq. (2·1). Regarding $F/(1+n)$ as the inverse temperature, this means that the evaluation of the average decoding error probability can be related to calculating the general moment of the 'partition function' $Z = \sum_{A=1}^{2^N-1}\exp\left[-\frac{F}{1+n}\epsilon_A\right]$ of the DREM.

Appendix B
—— *The Replica Analysis of the GCDREM* ——

As found in §3, the generalized moment of the partition function can be evaluated *without using the RM* for the GCDREM, even when the system size N is finite. However, for direct comparison to the conventional analysis, it may be helpful to demonstrate the conventional RM-based analysis as well. Therefore, here we provide a brief sketch of the replica calculation of the GCDREM.

From Eq. (3·1), we first obtain the expression

$$\langle Z^n\rangle = \omega^{-\frac{nM}{2}}\sum_{i_1=0}^{M}\sum_{i_2=0}^{M}\cdots\sum_{i_n=0}^{M}\left\langle\prod_{\mu=1}^{n}n_{i_\mu}\right\rangle\omega^{\sum_{\mu=1}^{n}i_\mu}$$
$$= \omega^{-\frac{nM}{2}}\sum_{i_1=0}^{M}\sum_{i_2=0}^{M}\cdots\sum_{i_n=0}^{M}\prod_{i=1}^{M}\left(\left\langle n_i^{\sum_{\mu=1}^{n}\delta_{i,i_\mu}}\right\rangle\omega^{i\sum_{\mu=1}^{n}\delta_{i,i_\mu}}\right), \qquad (\text{B·1})$$

for $n = 1, 2, \cdots$, where $\langle\cdots\rangle$ represents the average over the Poisson distribution, Eq. (3·2). Let us next evaluate the candidates for the most dominant contribution in the limit $N \to \infty$ under the RS1, RS2 and 1RSB assumptions, following the conventional scheme of the RM. We focus on the case $\alpha > 1$ in accordance with the calculation of the CDREM.

Before proceeding further, it may be worth mentioning that the dynamical variables in the GCDREM are not single states, but the energy levels $i = 0, 1, 2, \cdots, M$, each of which is composed of multiple states. This difference makes composition of the solutions slightly different from that in the CDREM, presented in §2, although the final result is the same.

- RS1
  In RS1, all n replica levels are assumed to be allocated to n different levels i ($= 0, 1, 2, \cdots, M$). Therefore, for a given level i, we have

$$\sum_{\mu=1}^{n}\delta_{i,i_\mu} = \begin{cases} 1 & \text{when } i \text{ is one of the } n \text{ allocated energy levels,} \\ 0 & \text{otherwise.} \end{cases} \qquad (\text{B·2})$$

When i is an allocated level, the relation

$$\left\langle n_i^{\sum_{\mu=1}^n \delta_{i,i_\mu}} \right\rangle = \gamma_i \rightarrow \exp\left[N\left((1-\alpha)\log 2 + \alpha H(x)\right)\right] \qquad (\text{B}\cdot 3)$$

and

$$\omega^{i\sum_{\mu=1}^n \delta_{i,i_\mu}} = \exp\left[-\beta i\right] \rightarrow \exp\left[-N\alpha\beta x\right] \qquad (\text{B}\cdot 4)$$

hold for large N. This yields the contribution

$$\left\langle n_i^{\sum_{\mu=1}^n \delta_{i,i_\mu}} \right\rangle \omega^{i\sum_{\mu=1}^n \delta_{i,i_\mu}} \rightarrow \exp\left[N\left((1-\alpha)\log 2 + \alpha(H(x)-\beta x)\right)\right], (\text{B}\cdot 5)$$

where $x \equiv i/M = i/(\alpha N)$. This takes its maximum,

$$\exp N\left(\log 2 + \alpha\log\left(\cosh\frac{\beta}{2}\right) - \frac{\alpha\beta}{2}\right), \qquad (\text{B}\cdot 6)$$

at $x_c = (1 - \tanh\frac{\beta}{2})$. This maximal value represents the contribution from a certain replica level i_μ $(= 1, 2, \cdots, n)$. Contributions from other replicas are smaller than that in Eq. (B·6), as any two replica levels must be located at different levels under the current assumption. However, the difference becomes negligible, because there exist many levels in any vicinity of $x = x_c$ in the limit $N \rightarrow \infty$. Taking the prefactor $\omega^{-\frac{nM}{2}}$ of the summation in Eq. (B·1) and the number of permutations over replica indices into account, this implies that the dominant contribution under the RS1 ansatz is given by

$$\langle Z^n \rangle = n! \times \exp nN\left(\log 2 + \alpha\log\left(\cosh\frac{\beta}{2}\right)\right)$$

$$\sim \exp nN\left(\log 2 + \alpha\log\left(\cosh\frac{\beta}{2}\right)\right), \qquad (\text{B}\cdot 7)$$

which is identical to the RS1 solution of the CDREM, Eq. (2·10), and is equivalent to Eq. (3·30).

• RS2

In RS2, all n replica levels are assumed to be allocated to a certain single level. Therefore, for a given level i, we have

$$\sum_{\mu=1}^n \delta_{i,i_\mu} = \left\{ \begin{array}{ll} n & \text{when } i \text{ is the allocated energy level,} \\ 0 & \text{otherwise.} \end{array} \right. \qquad (\text{B}\cdot 8)$$

When i is the allocated level, the relation

$$\left\langle n_i^{\sum_{\mu=1}^n \delta_{i,i_\mu}} \right\rangle \sim \left\{ \begin{array}{ll} \gamma_i^n \rightarrow \exp\left[nN\left((1-\alpha)\log 2 + \alpha H(x)\right)\right], & (x > x^*) \\ \gamma_i \rightarrow \exp\left[N\left((1-\alpha)\log 2 + \alpha H(x)\right)\right], & (x < x^*) \end{array} \right. (\text{B}\cdot 9)$$

where x^* is determined by Eq. (3·25), and

$$\omega^{i\sum_{\mu=1}^n \delta_{i,i_\mu}} = \exp\left[-n\beta i\right] \rightarrow \exp\left[-nN\alpha\beta x\right] \qquad (\text{B}\cdot 10)$$

hold for large N. Note that $\left\langle n_i^{\sum_{\mu=1}^{n} \delta_{i,i_\mu}} \right\rangle$ behaves differently for $x > x^*$ and $x < x^*$. Therefore, the behavior of the dominant contribution obtained by maximizing $\left\langle n_i^{\sum_{\mu=1}^{n} \delta_{i,i_\mu}} \right\rangle \omega^{i \sum_{\mu=1}^{n} \delta_{i,i_\mu}}$ depends the relation between x^* and the position of the maximum x_c. For $\beta < \beta_c$ and sufficiently small $n > 0$, $x_c = (1 - \tanh \frac{\beta}{2})/2$ becomes greater than x^*. This results in the RS1 solution, Eq. (2·10), in spite of the fact that the RS2 ansatz is presently employed. This is because the RS2 ansatz for *energy levels* does not necessarily imply that the n replica *states* are identical, in particular, when the occupation number n_i increases exponentially for large N, which is the case for $x_c > x^*$. In such cases, even if n replica states are allocated to the same energy level, they are typically distributed into n different states at the energy level. This corresponds to the RS1 ansatz of the CDREM, and, therefore, Eq. (2·10) should be reproduced. On the other hand, for $\beta > \beta_c$, the quantity $\left\langle n_i^{\sum_{\mu=1}^{n} \delta_{i,i_\mu}} \right\rangle \omega^{i \sum_{\mu=1}^{n} \delta_{i,i_\mu}}$ is maximized at $x_c = (1 - \tanh \frac{n\beta}{2})/2 < x^*$. This yields

$$\langle Z^n \rangle = \exp N \left(\log 2 + \alpha \log \left(\cosh \frac{n\beta}{2} \right) \right), \qquad \text{(B·11)}$$

which is identical to the RS2 solution of the CDREM, Eq. (2·13), and equivalent to Eq. (3·34). As Eq. (2·13) can be dominant at $n = 1, 2, \cdots$ only for $\beta > \beta_c$, this is consistent with the result of the replica analysis of the CDREM.

- 1RSB

In 1RSB, n replica levels are assumed to be equally allocated to n/m levels. Therefore, for a given level i, we have

$$\sum_{\mu=1}^{n} \delta_{i,i_\mu} = \begin{cases} m & \text{when } i \text{ is the } n/m \text{ allocated energy levels,} \\ 0 & \text{otherwise.} \end{cases} \qquad \text{(B·12)}$$

When i is the allocated level, the relation

$$\left\langle n_i^{\sum_{\mu=1}^{n} \delta_{i,i_\mu}} \right\rangle \sim \begin{cases} \gamma_i^m \to \exp\left[mN\left((1-\alpha)\log 2 + \alpha H(x)\right)\right], & (x > x^*) \\ \gamma_i \to \exp\left[N\left((1-\alpha)\log 2 + \alpha H(x)\right)\right], & (x < x^*) \end{cases} \qquad \text{(B·13)}$$

and

$$\omega^{i \sum_{\mu=1}^{n} \delta_{i,i_\mu}} = \exp\left[-m\beta i\right] \to \exp\left[-mN\alpha\beta x\right] \qquad \text{(B·14)}$$

hold for large N. Similarly to the case of RS2, this reproduces the RS1 solution, Eq. (2·10), for $\beta < \beta_c$ and sufficiently small $n > 0$. Therefore, we focus on the low temperature phase $\beta > \beta_c$. Then, $\left\langle n_i^{\sum_{\mu=1}^{n} \delta_{i,i_\mu}} \right\rangle \omega^{i \sum_{\mu=1}^{n} \delta_{i,i_\mu}}$ takes its maximum value,

$$\exp N \left(\log 2 + \alpha \log \left(\cosh \frac{m\beta}{2} \right) - \frac{m\alpha\beta}{2} \right), \qquad \text{(B·15)}$$

at $x_c = (1 - \tanh \frac{m\beta}{2})/2 < x^*$. This maximal value represents the contribution from one of the n/m allocated levels. The number of ways to select n/m from M

levels, equally allocating n replicas, is negligible in the subsequent calculation. Taking contributions from all n/m allocated levels and the extremization with respect to m into account, we obtain

$$
\begin{aligned}
\langle Z^n \rangle &\sim \exp N \left(\operatorname*{extr}_{m} \left\{ \frac{n}{m} \left(\log 2 + \alpha \log \left(\cosh \frac{m\beta}{2} \right) \right) \right\} \right) \\
&= \exp N n \alpha \beta \left(\frac{1}{2} - x^* \right) \\
&= \exp N \left(\frac{n \alpha \beta}{2} \tanh \frac{\beta_c}{2} \right),
\end{aligned}
\tag{B·16}
$$

which is identical to the 1RSB solution of the CDREM, Eq. (2·21), and is equivalent to Eq. (3·33). As Eq. (2·21) can be dominant at $n = 1, 2, \cdots$ only for $\beta > \beta_c$, this is consistent with the result of the replica analysis of the CDREM.

Appendix C
—— *Efficient Sampling in the CDREM* ——

Here, we show that $\{n_i\}$ can be sampled with an $O(M)$ computational cost in both the CDREM and the GCDREM. With the probability distribution for each level,

$$
P(E_i) = 2^{-M} \begin{pmatrix} M \\ \frac{1}{2}M + E_i \end{pmatrix}, \quad \left(E_i = i - \frac{M}{2} \right),
\tag{C·1}
$$

the probability to sample a configuration $(n_0, n_1, \cdots, n_M)$ is

$$
\mathcal{P}(n_0, n_1, \cdots, n_M) = \{P(E_0)\}^{n_0} \{P(E_1)\}^{n_1} \cdots \{P(E_M)\}^{n_M} \frac{2^N!}{n_0! n_1! \cdots n_M!}.
\tag{C·2}
$$

To generate a sample $(n_0, n_1, \cdots, n_M)$ in practice, we first determine n_0 according to the probability

$$
\mathcal{P}(n_0, \text{arbitrary}) = (p_0)^{n_0} (1 - p_0)^{2^N - n_0} \frac{2^N!}{n_0! (2^N - n_0)!}
\tag{C·3}
$$

$$
(p_0 \equiv P(E_0)).
\tag{C·4}
$$

We then determine n_1 according to the probability

$$
\mathcal{P}(n_0; n_1, \text{arbitrary}) = (p_1)^{n_1} (1 - p_1)^{2^N - n_0 - n_1} \frac{2^N!}{n_1! (2^N - n_0 - n_1)!}
\tag{C·5}
$$

$$
\left(p_1 \equiv \frac{P(E_1)}{1 - P(E_0)} \right).
\tag{C·6}
$$

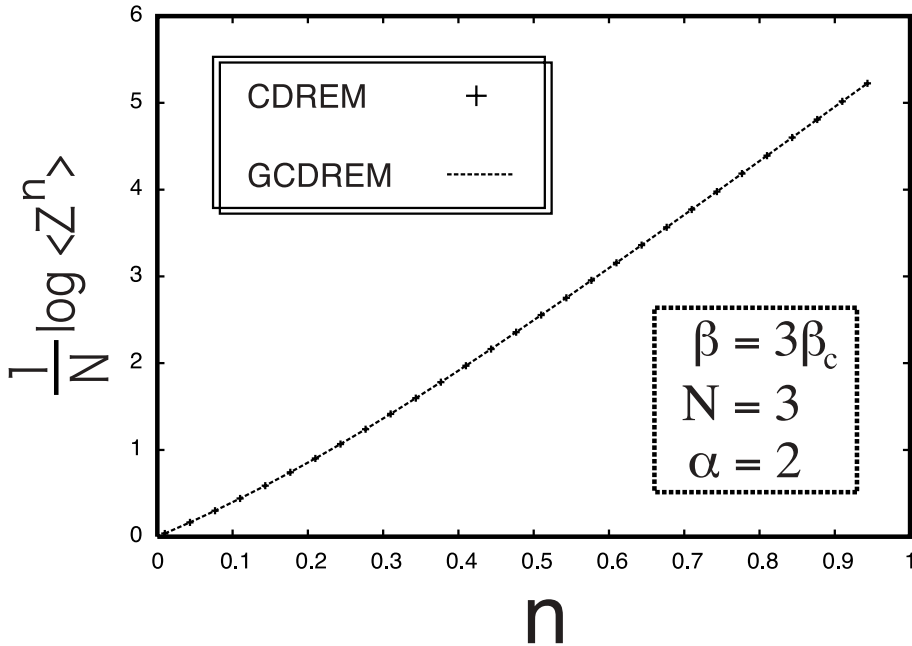Repeating this procedure up to n_{M-1}, we have

Fig. 11. The free energies obtained from the sets $\{n_i\}$ using the CDREM and the GCDREM. We find that the two models give the almost same results, even for $N = 3$.

$$\mathcal{P}(n_0, n_1, \cdots; n_{M-1}, n_M) = (p_{M-1})^{n_1}(1 - p_{M-1})^{2^N - \sum_{i=0}^{M-1} n_i} \frac{2^N!}{n_{M-1}!(2^N - \sum_{i=0}^{M-1} n_i)!}$$

$$\left(p_{M-1} \equiv \frac{P(E_{M-1})}{1 - \sum_{i=0}^{M-2} P(E_i)} \right) \tag{C·7}$$

and thereby obtain a set of $(n_0, n_1, \cdots, n_M)$. This guarantees that the identity

$$\sum_{i=0}^{M} n_i = 2^N \tag{C·8}$$

holds. This identity characterizes the CDREM. As in the case of the GCDREM, this can be performed in an $O(M)$ computation.

A comparison between the numerically evaluated moments of the partition functions for the CDREM and the GCDREM is presented in Fig. 11. It reveals that results of the two models are almost indistinguishable, even for $N = 3$. Because the consistency becomes better as N increases,[20),21)] using the GCDREM instead of the CDREM is justified when N is large.

References

1) M. Mézard, G. Parisi and M. A. Virasoro, *Spin glass theory and beyond* (World Scientific, Singapore, 1987).
2) S. F. Edwards and P.W. Anderson, J. of Phys. F **5** (1975), 965.
3) D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35** (1975), 1792.
4) G. Parisi, Phys. Lett. **73A** (1979), 203.
5) G. H. Hardy, Messenger Math. **58** (1929), 115.
6) F. Riesz, J. London Math. Soc. **5** (1930), 120.
7) G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities* (Cambridge University Press, Cambridge, 1934).
8) H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing* (Oxford University Press, New York, 2001).
9) N. Sourlas, Nature **339** (1989), 693.
10) Y. Kabashima and D. Saad, Europhys. Lett. **45** (1999), 97.
11) H. Nishimori and K. Y. M. Wong, Phys. Rev. E **60** (1999), 132.
12) K. Tanaka, J. of Phys. A **35** (2002), R81.
13) T. L. H. Watkin, A. Rau and M. Biehl, Rev. Mod. Phys. **65** (1993), 499.
14) R. Monasson and R. Zecchina, Phys. Rev. Lett. **76** (1996), 3881.
15) E. Korutcheva, M. Opper and L. López, J. of Phys. A **27** (1994), L645.
16) D. B. Saakian, cond-mat/0310549.
17) T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, Inc, New York, 1991).
18) V. Vapnik, *The Nature of Statistical Learning Theory* (Springer-Verlag, New York, 1995).
19) B. Derrida, Phys. Rev. B **24** (1981), 2613.
20) C. Moukarzel and N. Parga, Physica A **177** (1991), 24.
21) C. Moukarzel and N. Parga, Physica A **185** (1992), 305.
22) E. C. Titchmarsh, *The Theory of Functions* 2nd ed. (Oxford University Press, Oxford, 1939).
23) J. L. van Hemmen and R. G. Palmer, J. of Phys. A **12** (1979), 563.
24) T. Horiguchi, J. Math. Phys. **20** (1979), 1774.
25) C. E. Shannon, Bell. Syst. Tech. J. **27** (1948), 379; Bell. Syst. Tech. J. **27** (1948), 623.
26) Y. Kabashima, N. Sazuka, K. Nakamura and D. Saad, Phys. Rev. E **64** (2001), 046113.
27) Y. Kabashima, T. Murayama and D. Saad, Phys. Rev. Lett. **84** (2000), 1355.
28) N. Skanzos, J. van Mourik, D. Saad and Y. Kabashima, J. of Phys. A **36** (2003), 11131.
29) J-P. Bouchaud and M. Mezard, J. of Phys. A **30** (1997), 7997.
30) R. Monasson and D. O'Kane, Europhys. Lett. **27** (1994), 85.
31) R. W. Penney, T. Coolen and D. Sherrington, J. of Phys. A **26** (1993), 3681.
32) E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
33) C. N. Yang and T. D. Lee, Phys. Rev. **87** (1952), 404; Phys. Rev. **87** (1952), 410.
34) K. Ogure and Y. Kabashima, in preparation.
35) R. G. Gallager, *Information Theory and Reliable Communication* (Wiley, New York, 1968).
36) Y. Kabashima and D. Saad, J. of Phys. A **37** (2004), R1.
37) Y. Iba, *The Gallager formalism in information theory and the replica method* (in Japanese), (1989), unpublished note.
38) Y. Iba, J. of Phys. A **32** (1999), 3875.