Large Deviations in the Free Energy of Mean-Field Spin Glasses

Giorgio Parisi[1] and Tommaso Rizzo[2]

[1]*Dipartimento di Fisica, Università di Roma "La Sapienza", P.le Aldo Moro 2, 00185 Roma, Italy*
[2]*"E. Fermi" Center, Via Panisperna 89 A, Compendio Viminale, 00184, Roma, Italy*

(Received 8 June 2007; revised manuscript received 24 June 2008; published 11 September 2008)

We compute analytically the probability distribution of large deviations in the spin-glass free energy for the Sherrington-Kirkpatrick mean-field model; i.e., we compute the exponentially small probability of finding a system with intensive free energy smaller than the most likely one. This result is obtained by computing the average value of the partition function to the power n as a function of n. At zero temperature this absolute prediction displays a remarkable quantitative agreement with the numerical data.

DOI: 10.1103/PhysRevLett.101.117205 PACS numbers: 75.10.Nr

In the study of disordered systems nearly all predictions concern the most likely behavior, but there is also considerable interest in developing techniques to compute the probability distribution of rare events, i.e., the probability of finding systems that have properties different from the typical ones. The motivations are various. (i) We may have special interest in those systems with a behavior different from the most likely one; for example, in constraint optimization problems, in the region where it is impossible to satisfy all the constraints in the most likely system, there is a great interest in computing the properties of those rare systems where we can find a configuration that satisfies all the constraints [1]. (ii) The properties of large fluctuations may be related to other more interesting properties of the system. For example, given an intensive quantity A_J that depends on the system J of size N, in the large deviation region for large N we usually have that $P_N(A) \approx \exp(-NL(A))$. It is quite common that there are relations among the behavior of $P_N(A)$ in the region where the probability remains finite when N goes to infinity and the behavior of $L(A)$ near the point where $L(A) = 0$. In other cases [2], the techniques used to compute large deviations are the same used to compute more mundane quantities like (in finite dimensional spin glasses) the typical difference of the energy with periodic and antiperiodic boundary conditions. Besides, sample-to-sample fluctuations have been recently shown to be related to chaos in spin glasses [3]. (iii) We notice also that the comparison between analytic predictions in the large deviations region and numerical or experimental data may provide a clear-cut test of the theoretical approach used to compute the most likely properties.

Unfortunately even in the simplest nontrivial case, i.e., the Sherrington-Kirkpatrick (SK) infinite range model for spin glasses, there is no consensus on the results of such a computation. Everybody agrees that as a first step we need to compute in the large N limit the thermodynamic function

$$\Phi(n, \beta) = -\frac{1}{\beta n N} \ln \overline{Z_J(\beta)^n}, \tag{1}$$

where different systems (or samples) are labeled by J, $Z_J(\beta)$ is the partition function and the bar denotes the average over different disordered samples. It is well known that the probability of large deviations is related to the function $\Phi(n, \beta)$. Indeed,

$$\exp(\beta n N \Phi(n, \beta)) = \overline{Z_J(\beta)^n} = \overline{\exp(-nN\beta f_J(\beta))}, \tag{2}$$

where f_J is the system-dependent free energy *per spin*. The region of positive n corresponds to fluctuations where the free energy is smaller than the typical one and the region on negative n corresponds to fluctuations where the free energy is larger than the typical one.

There is a disagreement in the literature on the strategy we should follow to compute $\Phi(n, \beta)$. In the $n \to 0$ limit the computation can be done using the broken replica symmetry ansatz (that is known to give the exact result), where it coincides with the most likely free energy $\Phi(0, \beta) = f_{\rm typ}$ or, equivalently, with the average equilibrium free energy $\bar{f}_{\rm eq} = f_{\rm typ}$.

For $n > 0$ Kondor [4] in 1983 presented a first computation of $\Phi(n, \beta)$ in the region near T_c using the most natural ansatz for replica symmetry breaking (RSB) obtaining in the region of positive n

$$\Phi(n, \beta) = f_{\rm typ} + c_5 n^5 + O(n^6). \tag{3}$$

However, it has not been possible to test directly Kondor prediction because at the present moment all numerical data concern the fluctuations of the ground state energy; i.e., the system is at zero temperature. Indeed at zero temperature the free energy coincides with the internal energy and the numerical data are more clean due to the absence of thermal noise. The result of Kondor was surprising: in the general case all powers of n are present in the Taylor expansion of $\Phi(n, \beta)$ and for most of the systems we have $\Phi(n, \beta) = f_{\rm typ} + A_1 n + O(n^2)$, that is the typical situation for a Gaussian distribution of the free energy. The absence of the powers from n^1 to n^4 is due to cancellations and it was not clear if they were present only near the critical temperature. This form of the large deviation function implies that the probability

distribution for f near (and smaller than) $f_{\rm typ}$ is of the form

$$P_N(f) \propto \exp(-Na_{6/5}(f_{\rm typ}-f)^{6/5}), \tag{4}$$

where $a_{6/5} = 5\beta 6^{-6/5}|c_5|^{-1/5}$.

Many efforts has been concentrated on the scaling of the small deviations of the free energy. Indeed based on Kondor's result it was argued in [5] that the small deviations from its mean of the free energy per spin scale as $N^{-5/6}$. The function $P_N(f)$ is peaked around $f_{\rm typ}$ in the large N limit and the small deviations correspond to the region of the peak while the large deviations correspond to its exponentially small tails. Thus, in principle, small and large deviations are fairly different objects and it may seem strange that one can determine the scaling of the peak from a large deviation calculation. However, if one assumes that there is a smooth matching between the behavior of the peak and that of the tails, it follows that the region of the peak corresponds to values of the free energy difference such that expression (4) is finite and this leads to $(f_{\rm typ} - f) = O(N^{-5/6})$. This prediction has been tested in a series of numerical works [6–12] and although all estimates are smaller than $5/6$ nobody has claimed that this value is definitively ruled out. However, it was difficult to test the theory in the absence of a quantitative prediction (the only prediction being on the exponent, a quantity that it is rather difficult to measure in a reliable way). There are results that strongly indicate that the fluctuations of the internal energy per spin at finite temperature scale as $N^{-5/6}$, thus confirming Kondor's exponent [13].

More recently, a different RSB ansatz was proposed by Aspelmeier and Moore [14,15], who found $\Phi(n) = f_{\rm typ}$; in their approach the probability of large deviations goes to zero faster than $\exp(-L(f)N)$ and the above matching argument cannot be used to infer the small deviations exponent. Indeed there is a general agreement that for negative n $\Phi(n) = f_{\rm typ}$ and $P_N(f)$ goes to zero faster than $\exp(-CN)$ as soon $f > f_{\rm typ}$. It is quite possible that in that region we have $P_N(f) \propto \exp(-N^2 L_2(f))$ with some unknown function $L_2(f)$ as it happens in the spherical model [16]; however, an analysis of this point goes beyond the aim of this Letter.

In the following we concentrate on large deviations in the region $f < f_{\rm typ}$, that corresponds to positive n. We follow Kondor's approach and we extend his computation to all temperatures, including $T = 0$; in this way we obtain an absolute prediction for the large deviations distribution. We compare our analytic results with the numerical simulations performed at zero temperature and we find a remarkable agreement. We also find that the alternative approach [14,15] that predicts $\Phi(n) = f_{\rm typ}$ for both negative and positive values of n cannot be valid for large positive n and there are no compelling reasons for which it should be valid at fixed positive n when N goes to infinity. This is in agreement with the results coming from an exact analysis: for positive values of n Talagrand [17] was able to show rigorously that Kondor's approach gives the correct results. The solution to the problem of computing the large deviations for the SK model at all temperatures is presented below: eventually we will concentrate on the zero-temperature limit.

We start our analysis by defining carefully the large deviation function for the free energy, $L(f)$, (that we will call in the following the sample complexity because it is related to the number of samples with free energy equal to f) as the logarithm divided by N of the probability density of finding a sample of size N with free energy per spin f in the thermodynamic limit [5], i.e.,

$$L(f) = \lim_{N\to\infty} \frac{\log(P_N(f))}{N}. \tag{5}$$

For large N the majority of the samples has free energy per spin equal to $f_{\rm typ}$, and all other values have exponentially small probability. Consistently $L(f)$ is less or equal than zero, the equality holding for $f = f_{\rm typ}$, i.e., $L(f_{\rm typ}) = 0$. For some values of f it is possible that $L(f) = -\infty$, signaling that the probability of large deviations goes to zero faster than exponentially with N.

It is evident that $\Phi(n)$ is the Legendre transform of $L(f)$:

$$-\beta n\Phi(n) = -\beta n f + L(f), \qquad \beta n = \frac{\partial L}{\partial f}. \tag{6}$$

Equivalently, we have

$$L(f) = \beta n f - \beta n \Phi(n), \qquad f = \frac{\partial(n\Phi(n))}{\partial n}. \tag{7}$$

In the Sherrington-Kirkpatrick model at low temperatures the replica symmetry is spontaneously broken for the generic system, i.e., in the $n \to 0$ limit. One knows that at high positive values of n, replica symmetry is not broken [18]. Therefore for positive n one must distinguish two regions in the $T - n$ plane separated by the so called de Almeida–Thouless (dAT) line; see Fig. 1. In the region

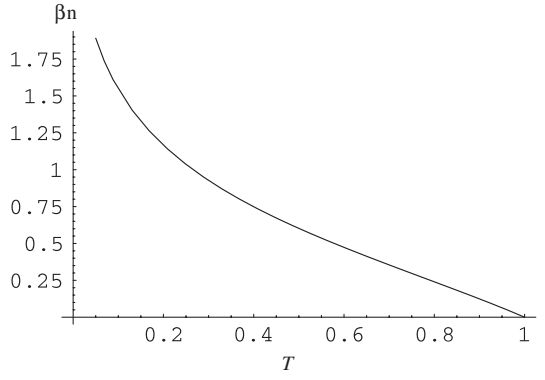

FIG. 1. The dAT line in the $(T, \beta n)$ plane. The value of βn diverges in the zero-temperature limit as $\beta n \simeq \sqrt{-2\ln T}$, as a consequence the function $L(\Delta e)$ at zero temperature is described by the RSB solution at any value of Δe.

above the dAT line, the phase is replica-symmetric, while replica symmetry is broken below.

In the replica-symmetric (RS) region the order parameter is the overlap q. The corresponding value of the potential $\Phi(n, q)$ is given by

$$\Phi(n, q) = -\frac{\beta}{4}(1 - 2q + (1 - n)q^2) + -\frac{1}{\beta n} \times \ln \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi q}} e^{-(y^2/2q)} (2\cosh \beta y)^n.$$

The overlap q can be computed by solving the equation $\partial\Phi(n, q)/\partial q = 0$. In the (n, T) plane the dAT line is specified by the condition [19]:

$$T^2 = \int dP(y)(1 - \tanh^2 \beta y)^2 dy, \tag{8}$$

where $P(y)$ is a normalized probability proportional to $e^{-(y^2/2q)}(\cosh\beta y)^n$.

On the dAT line the value of n is $n_{\mathrm{dAT}}(T) = 4\tau/3$ for small $\tau = 1 - T$ while $n_{\mathrm{dAT}}(T)$ vanishes in the zero-temperature limit as $n = T\sqrt{-2\ln T}$. As a consequences in the rescaled $(T, n\beta)$ plane the dAT line never touches the $T = 0$ line and $L(e)$ at $T = 0$ is always in the RSB phase; see Fig. 1.

For small n the RS solution is not only unstable but also inconsistent, indeed near the critical temperature (for $T < T_c$) one finds an unphysical positive value of the sample complexity. At any finite temperature $\Phi(n)$ is described by the RS solution at large values of n. Both above and below the critical temperature, the behavior of $\Phi(n)$ for large values of n is $\Phi(n) = -\beta n/4 - \ln 2/(\beta n) + O(e^{-2\beta n})$. This leads to $L(f) = -f^2 + \ln 2 + o(1)$ for large negative f, (this is the same behavior of the random energy model [20]).

Below the dAT line we must break the replica symmetry. An explicit computation shows that the free energy on the dAT line is smaller than the most likely free energy (f_{typ}), in particular, it diverges as $-\sqrt{-(\ln T)/2}$ at low temperatures while for small $\tau = T_c - T$ the free energy difference is $f_{\mathrm{eq}} - f_{\mathrm{dAT}} = 2\tau^5/45 + O(\tau^6)$. Therefore, we must look for a free energy that shows some dependence on n also below the dAT line and the one suggested by Kondor is the most natural one.

We recall that in Kondor's approach for $n < n_{\mathrm{dAT}}(T) < 1$, one introduces a function $q(x)$ defined for $n \le x \le 1$ that describes the breaking of replica symmetry in the low-temperature phase. A functional $F_n[q]$ is obtained such that $\Phi(n) = \max_q F_n[q]$. The function $q(x)$ that maximizes $F_n[q]$ can be found by solving the stationarity equation $\delta F/\delta q(x) = 0$. This generalizes the standard approach that is proved to give the correct value of $\Phi(n)$ in the $n \to 0$ limit.

The form of the free energy functional is the usual one [19], the only difference being that all functions are defined in the interval $n \le x \le 1$. Kondor [4] found that near the critical temperature Eq. (3) holds with $c_5 = -9/5120$. For negative n the saddle point of the $F_n[q]$ is the standard $q(x)$ corresponding to $n \to 0$, thus $\Phi(n) = f_{\mathrm{typ}}$ for $n < 0$ [21]. Near the critical temperature the corresponding sample complexity as a function of $\Delta f = f - f_{\mathrm{typ}}$ reads

$$L(f) = -\infty \quad \text{for } \Delta f > 0,$$
$$L(f) = -a_{6/5}|\Delta f|^{6/5} + O(|\Delta f|^{8/5}) \quad \text{for } \Delta f \le 0,$$

where, as we have already noticed, $a_{6/5} = 5\beta|c_5|^{-1/5}6^{-6/5}$.

There are many ways in which one can compute the maximum of $F_n[q]$ in a systematic way far away of the critical temperature. We have followed the strategy discussed in [22,23]. We have solved the RSB equations and computed $q(n, x)$ and $\Phi(n)$ as a series in powers of n and $\tau = 1 - T$ [22]. The computation showed that at all the orders considered that the lowest power of n in the expansion of $\Phi(n)$ is always n^5 and also that there is no n^6 term. The computation has been done up to the order τ^{18}.

We have verified by an expansion in powers of n at fixed τ that the first term in $\Phi(n)$ is of $O(n^5)$ at all temperatures as follows from an analytic argument that for reasons of space will be reported elsewhere [24]. This result is related to the behavior of the free energy functional with increasing number of RSB steps [25].

It is interesting to note that from the third order on, all derivatives of $\Phi(n)$ (with respect to n, T and both) are discontinuous on the dAT line; i.e., the transition is third order, using the old fashioned thermodynamic classification. This is the same behavior of the free energy on the dAT line in the (h, T) plane [26].

When $\beta \to \infty$ the sample complexity $L(f)$ goes to a well-defined limit. Therefore, from Eq. (6) $\Phi(n)$ at low temperature should behave as a function of βn [27] and the coefficient c_a of n^a in the power series of $\Phi(n)$ diverges as β^a in the zero-temperature limit.

The series in powers of τ of c_5 (the n^5 coefficient in $\Phi(n)$) can be used to obtain its behavior in the whole low-temperature phase provided one uses the information that $c_5 \sim \beta^5$ in the zero-temperature limit. Indeed, the series can be resummed using Padé approximants with estimated errors not greater that 1% in the whole temperature range. A similar result can be obtained for c_7.

Using the Padé approximants we can safely make an extrapolation to zero temperature. At $T = 0$ we find

$$L(\Delta e) = -1.62(1)|\Delta e|^{6/5} + 3.1(1)|\Delta e|^{8/5} + O(\Delta e^{8/5}). \tag{9}$$

Unfortunately, the second term yields a big correction to the first one and actually we expect the series to be asymptotic as is usually the case in this context [22].

In order to bypass this problem and to have a good control on $L(\Delta e)$ we have adopted a method introduced

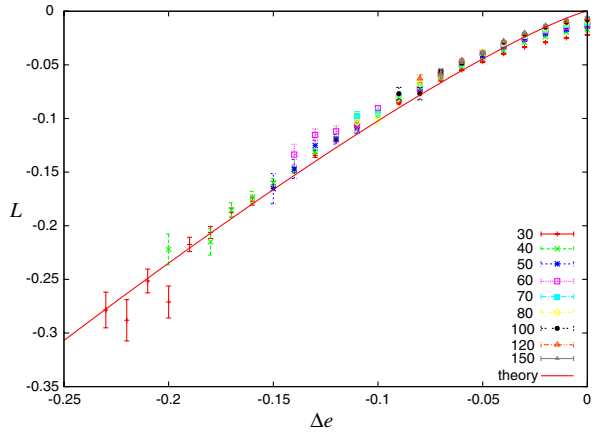

FIG. 2 (color online). Comparison between the numerical and analytical sample complexity at zero temperature, $L(\Delta e)$ versus Δe. The numerical (from $N = 30$ to $N = 150$) data are those of Ref. [8].

in [22] to obtain $q(x, \tau)$ from its series in powers of x and τ. We have transformed the series of $L(\Delta f)$ in powers of Δf and τ in a power series of just τ by setting $\Delta f = z(\frac{2}{45}\tau^5 + \frac{1}{4}\tau^7)$ with z a parameter. The corresponding series in powers of τ were resummed for any given z through Padé approximants obtaining in the low-temperature limit the curve $L(\Delta e)$ in parametric form. By resumming the series of $\Phi(n, \tau)$ as a function of τ we have been able to obtain the sample complexity in the whole low-temperature phase using the technique of Padé approximants: 18 orders of the Taylor expansion give us a very good control on the function.

Using this technique we find that, for not too negative Δe, the zero-temperature result shown in Fig. 2 differs by less than 1% from the first term of Eq. (9) in this range of energy differences. We compared the sample complexity with the numerical data at zero temperature of Ref. [8] as a function Δe and find a very good agreement. For each N we have plotted $L_N = \ln(P(\Delta e_N)/N^{5/6})/N$ with $\Delta e_N = e - e_N$ (the average energy at size N): we put the $N^{5/6}$ factor in the definition so that L_N should go to a constant for $\Delta e_N = 0$. The quantitative agreement of the numerical with the theory is quite good. A similar good agreement is asymptotically obtained if we plot the data at fixed energy.

Summarizing, using the standard hierarchical ansatz we have computed the large deviation function at all temperatures. In this way we have been able to confirm that the sample complexity $L(\Delta f)$ is proportional to $|\Delta f|^{6/5}$ for small negative Δf, this result strongly suggests that the sample-to-sample fluctuations are proportional to $N^{-5/6}$. We have verified that the numerical data of [8] are in remarkably good agreement with our absolute prediction. We believe that our results solve the problem of computing the large deviations function for negative Δf. One could therefore start to study more difficult problems, like the large deviation function for positive Δf in the SK model. One could also try to extend our results to other models, such as Bethe lattices or large dimensional short range models; work is in progress in these directions [24].

We thank the authors of Ref. [8] for giving us their numerical data.

[1] F. Altarelli, R. Monasson, and F. Zamponi, J. Phys. A **40**, 867 (2007).
[2] T. Aspelmeier, M. A. Moore, and A. P. Young, Phys. Rev. Lett. **90**, 127202 (2003).
[3] T. Aspelmeier, Phys. Rev. Lett. **100**, 117205 (2008).
[4] I. Kondor, J. Phys. A **16**, L127 (1983).
[5] A. Crisanti, G. Paladin, H.-J. Sommers, and A. Vulpiani, J. Phys. I (France) **2**, 1325 (1992).
[6] S. Boettcher, Europhys. Lett. **67**, 453 (2004).
[7] J.-P. Bouchaud, F. Krzakala, and O. C. Martin, Phys. Rev. B **68**, 224404 (2003).
[8] A. Andreanov, F. Barbieri, and O. C. Martin, Eur. Phys. J. B **41**, 365 (2004).
[9] M. Palassini, arXiv:cond-mat/0307713.
[10] S. Boettcher, Eur. Phys. J. B **46**, 501 (2005).
[11] H. G. Katzgraber, M. Korner, F. Liers, M. Junger, and A. K. Hartmann, Phys. Rev. B **72**, 094421 (2005).
[12] K. F. Pal, Physica (Amsterdam) **367A**, 261 (2006).
[13] T. Aspelmeier, A. Billoire, E. Marinari, and M. A. Moore, J. Phys. A (to be published).
[14] T. Aspelmeier and M. A. Moore, Phys. Rev. Lett. **90**, 177201 (2003).
[15] C. De Dominicis and P. Di Francesco, arXiv:cond-mat/0301066.
[16] D. S. Dean and S. N. Majumdar, Phys. Rev. Lett. **97**, 160201 (2006).
[17] M. Talagrand Large Deviations, Guerra's and A.S.S. Schemes, and the Parisi Hypothesis, to appear in The Proceedings of the Conference Mathematical Physics of Spin-Glasses, Cortona (2005).
[18] D. Sherrington, J. Phys. A **13**, 637 (1980).
[19] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[20] B. Derrida, Phys. Rev. B **24**, 2613 (1981).
[21] Vik. Dotsenko, S. Franz, and M. Mezard, J. Phys. A **27**, 2351 (1994).
[22] A. Crisanti and T. Rizzo, Phys. Rev. E **65**, 046137 (2002).
[23] H. J. Sommers and W. Dupont, J. Phys. C **17**, 5785 (1984).
[24] G. Parisi and T. Rizzo (to be published).
[25] R. Oppermann, M. J. Schmidt, and D. Sherrington, Phys. Rev. Lett. **98**, 127201 (2007).
[26] A. Crisanti, T. Rizzo, and T. Temesvari, Eur. Phys. J. B **33**, 203 (2003).
[27] This result is related to the scaling in the $T \to 0$ limit of the function $q(x)$: $q(x, \beta) \approx \hat{q}(\beta x)$.