The distribution of the partition function of the Hopfield model with finite number of patterns

1996 J. Phys. A: Math. Gen. 29 3911

(http://iopscience.iop.org/0305-4470/29/14/016)

J. Phys. A: Math. Gen. **29** (1996) 3911–3922. Printed in the UK

The distribution of the partition function of the Hopfield model with finite number of patterns

A E Patrick†

Limburgs Universitair Centrum, Departement WNI, Universitaire Campus, B-3590 Diepenbeek, Belgium

Received 29 September 1995, in final form 28 March 1996

Abstract. We derive the leading term in the large-N asymptotic expansion of the partition function of the Hopfield model with finite number of patterns. We show that this leading-order term is deterministic in the high-temperature region. In the low-temperature region and at the critical point it is random with the distribution governed by χ^2, normal, or iterated exponential distributions.

1. Introduction

Disordered systems of statistical mechanics are often defined in terms of random Hamiltonians $H_N(\boldsymbol{s}_N; \boldsymbol{\xi}_K(\omega))$, where $\boldsymbol{s}_N = \{s_i\}_{i=1}^N$ is a configuration of spins and $\boldsymbol{\xi}_K(\omega) \equiv \{\xi_i(\omega)\}_{i\in K}$ is a set of random parameters defined on some probability space $(\Omega, \mathcal{F}, \boldsymbol{P})$. The corresponding partition functions

$$Z_N \equiv \sum_{s_1,s_2,\ldots,s_N} \exp[-\beta H_N(\boldsymbol{s}_N; \boldsymbol{\xi}_K(\omega))]$$

are random functions with (as a rule) non-degenerate distributions. The free energy per spin

$$F_N(\beta) = -\frac{1}{\beta N} \log Z_N$$

is a random function with a nondegenerate distribution as well. However, as $N \to \infty$ the sequence $F_N(\beta)$ (as a rule) exhibits a self-averaging property, that is, the sequence of the corresponding distributions converges to a degenerate one in the thermodynamic limit [3, 8, 9, 12]. In this sense the limiting free energy $f(\beta) = \lim_{N\to\infty} F_N(\beta)$ is non-random. Moreover, the limiting free energy and the thermodynamic limit of the mean free energy coincide

$$f(\beta) = -\lim_{N\to\infty} \frac{1}{\beta N} \boldsymbol{E}(\log Z_N).$$

All these facts may create the impression that the distributions of (finite volume) partition functions are of no significance for the thermodynamic properties of disordered systems, especially due to the (unfortunately) rather widespread belief that all thermodynamic observables can be derived from the free energy by differentiation. However, as suggested by exactly solvable models (simplest mean field models [2] and disordered spherical models),

† Present address: Dublin Institute for Advanced Studies, School of Theoretical Physics, 10 Burlington Road, Dublin 4, Ireland. E-mail address: patrick@stp.dias.ie

0305-4470/96/143911+12$19.50

certain thermodynamic observables, like the magnetization, do not possess the self-averaging property in multiphase regions of phase diagrams. Therefore, knowledge of the limiting free energy is not sufficient for calculation of thermodynamic observables, and the distribution of the partition function is of significant importance for the derivation of distributions of non-self-averaging observables [2].

Another reason to study the distributions of random partition functions is the widespread application of the replica method. The starting point of the replica method is an attempt to obtain a generic expression for the integer moments of a partition function. If the generic integer moment is successfully computed, the noninteger moments obtained by an analytical continuation can provide very useful information. Knowledge of some general properties of the distribution of the partition function may help to find the correct analytical continuation. Unfortunately, as a rule, it is impossible to obtain an explicit expression for the generic integer moment of the partition function. However, calculation of the first few integer moments of the partition function is often possible and still may provide useful information on the thermodynamic properties of the model if we know in advance how randomness enters the expression for the partition function and if we have control over the corresponding large deviation probabilities.

Guided by analogy with finite-size scaling theory one can guess the following asymptotic expression for the partition function

$$Z_N = \exp(N f_N(\beta) + N^\rho r_N) \tag{1}$$

where $f_N(\beta)$ is the 'non-random' contribution to the free energy, $\lim_{N\to\infty} f_N(\beta) = f(\beta)$; $\rho < 1$, and r_N is a sequence of random variables (the random part of the free energy) such that $\lim_{N\to\infty} r_N \stackrel{d}{\to} r$, where r is a random variable with a proper, non-degenerate distribution. Of course, the N^ρ term may be distorted by $\log N$-type corrections. The exponent ρ in equation (1) is expected to be the same for large classes of models. Also, the limiting random variable r should have its distribution confined to a rather restricted class.

In the present paper we calculate the distribution of the partition function corresponding to the Hamiltonian

$$H_N = -\frac{1}{2N} \sum_{i\neq j}^{N} \sum_{p=1}^{M} \xi_i^{(p)} \xi_j^{(p)} s_i s_j - \varepsilon \sum_{i=1}^{N} \xi_i^{(q)} s_i \tag{2}$$

where $\{\xi_i^{(p)} = \pm 1\}_{i=1,\,p=1}^{N\;\;\,M}$ is a sequence of independent and identically distributed random variables (i.i.d.r.v.) with the shared distribution $\Pr[\xi_i^{(p)} = \pm 1] = \frac{1}{2}$, the variables $\{s_i = \pm 1\}_{i=1}^N$ are Ising spins, and M is kept fixed in the thermodynamic limit ($N \to \infty$).

The Hamiltonian (2) was used (apparently) for the first time by Pastur and Figotin [8] to construct an exactly solvable model of a spin glass. They found an exact expression for the free energy corresponding to that Hamiltonian in the thermodynamic limit using the approximating Hamiltonian method. Later on the Hamiltonian (2) was used by Hopfield [6] in a model of a neural network exhibiting associative memory which since then has attracted much attention especially in the case when M grows linearly with N. In the framework of the Hopfield model the subsequences $\{\xi_i^{(p)}\}_{i=1}^N$, $p = 1, 2, \ldots, M$ represent patterns which are supposed to be stored by the network.

The term $-\varepsilon \sum_{i=1}^N \xi_i^{(q)} s_i$ in equation (2) corresponds to an external field favouring one of the equilibrium states (in the spin-glass terminology) or a nominated pattern (in the neural network interpretation). Such a field, however, is usually considered as rather artificial. We use it mainly in order to take the limit $\varepsilon \to 0$ after $N \to \infty$. This procedure enables one to obtain (possibly metastable) states (that is, probability measures) corresponding to an intermediate stage of the time evolution ($t \gg 1$, but $t \ll \tau^*$, where τ^* is a characteristic time of intervalley transitions for, for example, Glauber dynamics) generated by the Hamiltonian (2) with $\varepsilon = 0$. The states obtained in the thermodynamic limit with ε set to 0 beforehand correspond to a rough time-scale description of the evolution ($t \gg \tau^*$). Although for mean-field models, like (2), the characteristic time τ^* is very large (grows exponentially with N) consideration of the time-scales $t \gg \tau^*$ is not totally pointless. Indeed, for finite-dimensional models the height of the intervalley barriers is known to have a behaviour qualitatively different from the mean-field one.

It is appropriate to point out here that the overlap parameters

$$m_N^{(p)} = \frac{1}{N}\sum_{j=1}^{N} \xi_j^{(p)} s_j \qquad p = 1, 2, \ldots, M$$

in the Hopfield model, *even* with a finite number of patterns (the Hamiltonian (2) with $\varepsilon = 0$), are non-self-averaging observables. Like the magnetization of the Curie–Weiss model in random field [2] the overlap parameters are random variables with non-degenerate distributions even in the limit $N \to \infty$. In fact, $\lim_{N\to\infty} m_N^{(p)}$ exists only in distribution. Therefore, for a fixed realization of patterns $\{\xi_i^{(p)}\}_{i=1}^{\infty}$, $p = 1, 2, \ldots, M$ the limit $\lim_{N\to\infty} m_N^{(p)}$ simply does not exists (with probability one). As a consequence, any approach based on a reckless averaging is bound to produce, strictly speaking, wrong results for the Hopfield and similar models. The best one can hope for when applying, for example, the replica method to the model (2) in the case $\varepsilon = 0$ is to obtain, say, for the overlap parameter $m^{(q)}$ the results corresponding to taking limit $\varepsilon \downarrow 0$ (after the limit $N \to \infty$) in the expression for the overlap parameter $m^{(q)}(\varepsilon)$. Indeed, the overlap parameter $m_N^{(q)}(\varepsilon)$ in the model (2) is a self-averaging observable for any $\varepsilon \neq 0$. The results of the papers [9, 10] suggest, however, that in the case of the Hopfield model with an extensively large number of patterns it is very unreasonable to expect restoration of self-averaging by the field $\varepsilon \sum_{j=1}^{N} \xi_j^{(q)} s_j$ with arbitrarily small ε.

In order to keep the size of the paper small and to avoid repetition of (by now) well known arguments the author did not make an attempt to write a rigorous paper. However, the necessary polishing could have been done (by losing the advantages of brevity) and the interested readers can get an idea of how the missing proofs can be accomplished from [4] and references therein.

2. Distribution of the partition function

We consider first the Hamiltonian (2) with $\varepsilon = 0$. The partition function of the model is then given by

$$\begin{aligned} Z_N &= \sum_{s_1,\ldots,s_N} \exp\left[\frac{\beta}{N}\sum_{i<j}^{N}\sum_{p=1}^{M} \xi_i^{(p)}\xi_j^{(p)} s_i s_j\right] \\ &= \mathrm{e}^{-\frac{1}{2}M\beta} \sum_{s_1,\ldots,s_N} \exp\left[\frac{\beta}{2N}\sum_{p=1}^{M}\left(\sum_{i=1}^{N}\xi_i^{(p)} s_i\right)^2\right]. \end{aligned} \tag{3}$$

Using M times the well known identity

$$\mathrm{e}^{\frac{1}{2}ay^2} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-\frac{1}{2}x^2 + \sqrt{a}yx} \tag{4}$$

equation (3) can be rewritten in a form convenient for application of the Laplace method

$$Z_N e^{\frac{1}{2}M\beta} = \sum_{s_1,\ldots,s_N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{k=1}^{M} \frac{dx_k}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\sum_{p=1}^{M} x_p^2 + \sqrt{\frac{\beta}{N}} \sum_{p=1}^{M} x_p \sum_{i=1}^{N} \xi_i^{(p)} s_i\right]$$
$$= 2^N \left(\frac{\beta N}{2\pi}\right)^{M/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{k=1}^{M} dx_k\, e^{-N\Phi_N(x_1,x_2,\ldots,x_M)} \tag{5}$$

where

$$\Phi_N(\boldsymbol{x}) \equiv \Phi_N(x_1,\ldots,x_M) = \frac{1}{2}\beta\sum_{p=1}^{M} x_p^2 - \frac{1}{N}\sum_{i=1}^{N} \log\cosh\left(\beta\sum_{p=1}^{M} x_p \xi_i^{(p)}\right). \tag{6}$$

To evaluate the multiple integral in (5) using the Laplace method one has to find the dominant minima of the function $\Phi_N(\boldsymbol{x})$. As usual it is convenient to consider the cases $\beta < \beta_c$ (the high-temperature region), $\beta > \beta_c$ (the low-temperature region), and $\beta = \beta_c$ separately ($\beta_c = 1$ for the Hopfield model with finite number of patterns).

2.1. The low-temperature region

Consider the function

$$\Phi(\boldsymbol{x}) \equiv \Phi(x_1,\ldots,x_M) = \frac{1}{2}\beta\sum_{p=1}^{M} x_p^2 - \frac{1}{2^M}\sum_{\sigma_1,\ldots,\sigma_M=\pm1} \log\cosh\left(\beta\sum_{p=1}^{M}\sigma_p x_p\right) \tag{7}$$

and note that $\Phi_N(\boldsymbol{x}) \to \Phi(\boldsymbol{x})$ as $N \to \infty$ (with probability one). The function $\Phi(\boldsymbol{x})$ attains its global minimum at the $2M$ points $\pm\boldsymbol{x}_p^* = \{\pm x^*\delta_{p,k}\}_{k=1}^{M}$, $p = 1, 2, \ldots, M$; where x^* is a minimum point of the function

$$\Phi(x) = \tfrac{1}{2}\beta x^2 - \log\cosh\beta x$$

and $\delta_{p,k}$ is the Kronecker delta. The stationary points of the function $\Phi(x)$ are solutions of $x = \tanh\beta x$, among which the minimum points can be identified as the stable fixed points of the recurrent relation

$$x(n+1) = \tanh\beta x(n) \tag{8}$$

which is simply an algorithm for solving the the equation $x = \tanh\beta x$ by iterations.

The stationary points of the function $\Phi_N(x_1,\ldots,x_M)$ are the solutions of the system

$$x_p = \frac{1}{N}\sum_{i=1}^{N}\xi_i^{(p)}\tanh\left(\beta\sum_{f=1}^{M}x_f\xi_i^{(f)}\right) \qquad p = 1, 2, \ldots, M. \tag{9}$$

The minima of $\Phi_N(x_1,\ldots,x_M)$ are the stable fixed points of the system of recurrent relations

$$x_p(n+1) = \frac{1}{N}\sum_{i=1}^{N}\xi_i^{(p)}\tanh\left(\beta\sum_{f=1}^{M}x_f(n)\xi_i^{(f)}\right) \qquad p = 1, 2, \ldots, M \tag{10}$$

which, similar to (8), is an algorithm for solving the system (9) by iterations.

In general, it is not an easy problem to find the initial conditions for the recurrent relations (10) which are in the basins of attraction of the fixed points corresponding to the dominant minima of the function $\Phi_N(\boldsymbol{x})$. Fortunately, in our case (finite M) the choice is easy. Consider $2M$ balls $S(\pm\boldsymbol{x}_p^*) \subset \boldsymbol{R}^M$ centred at $\pm\boldsymbol{x}_p^*$, $p = 1, 2, \ldots, M$, which have radius r (independent of N) sufficiently small to contain only one point of the minima of the function $\Phi(\boldsymbol{x})$—the one in the centre of a ball. If N is large enough the system

(9) has exactly one solution $\boldsymbol{x}^*_{p,N} = (x^{(p)}_{1,N}, x^{(p)}_{2,N}, \ldots, x^{(p)}_{M,N})$ in each of the balls $S(\boldsymbol{x}^*_p)$ and $\lim_{N\to\infty} \boldsymbol{x}^*_{p,N} = \boldsymbol{x}^*_p$. This statement is proved sometimes by reference to the generalization of the Hurwitz theorem to functions of M complex variables. A more direct proof, based on the observation that the first and second partial derivatives of the function $\Phi_N(\boldsymbol{x})$ converge to those of $\Phi(\boldsymbol{x})$, is also possible [2]. Thus one can take as the initial conditions for the system of recurrent relations (10) the points $\pm\boldsymbol{x}^*_p$, $p = 1, 2, \ldots, M$, then solving the recurrent relation (10) will provide us with $2M$ fixed points $\pm\boldsymbol{x}^*_{p,N}$, $p = 1, 2, \ldots, M$. The only point which will need to be checked after that, to make sure that we have found the right minima, is that $\lim_{N\to\infty} \boldsymbol{x}^*_{p,N} = \boldsymbol{x}^*_p$. Typically, for a given N, the points of only one pair $\pm\boldsymbol{x}^*_{p,N}$ correspond to the global minima of the function $\Phi_N(\boldsymbol{x})$. Nevertheless, all the points $\pm\boldsymbol{x}^*_{p,N}$, $p = 1, 2, \ldots, M$ contribute to the main asymptotics of the integral in (5) and any other minima of the function $\Phi_N(\boldsymbol{x})$ are irrelevant.

Choosing $(x_1(0), x_2(0), \ldots, x_M(0)) = \boldsymbol{x}^*_q$ one obtains

$$x_p(1) = \frac{1}{N}\sum_{j=1}^{N} \xi^{(q)}_j \xi^{(p)}_j x_*$$

for $p = 1, 2, \ldots, M$. The second iteration yields

$$\begin{aligned} x_p(2) &= \frac{1}{N}\sum_{j=1}^{N} \xi^{(p)}_j \tanh\left(\beta \sum_{f=1}^{M} x_f(1)\xi^{(f)}_j\right) \\ &= \tanh[\beta x_q(1)]\frac{1}{N}\sum_{j=1}^{N}\xi^{(q)}_j\xi^{(p)}_j + \frac{\beta}{\cosh^2[\beta x_q(1)]}\sum_{f(\neq q)} x_f(1)\frac{1}{N}\sum_{j=1}^{N}\xi^{(f)}_j\xi^{(p)}_j \\ &\quad + \cdots \end{aligned}$$

where we have expanded $\tanh(x)$ in the Taylor series around the point $x = \beta x_q(1)$. Thus

$$x_p(2) = x_*\frac{1}{N}\sum_{j=1}^{N}\xi^{(q)}_j\xi^{(p)}_j + \frac{\beta}{\cosh^2(\beta x_*)}x_p(1) + \mathrm{O}(N^{-1})$$

for $p \neq q$, and

$$x_q(2) = \tanh[\beta x_q(1)] + \frac{\beta}{\cosh^2[\beta x_q(1)]}\sum_{f(\neq q)}^{M} x_f(1)\frac{1}{N}\sum_{j=1}^{N}\xi^{(f)}_j\xi^{(q)}_j + \mathrm{O}(N^{-3/2}).$$

On making further iterations in (10) it becomes clear that the generic expressions for $x_p(n)$ are given by

$$x_p(n+1) = x_*\frac{1}{N}\sum_{j=1}^{N}\xi^{(q)}_j\xi^{(p)}_j + \frac{\beta}{\cosh^2\beta x_*}x_p(n) + \mathrm{O}(N^{-1}) \tag{11}$$

if $p \neq q$, and

$$x_q(n+1) = \tanh[\beta x_q(n)] + \frac{\beta}{\cosh^2[\beta x_q(n)]}\sum_{f(\neq q)} x_f(n)\frac{1}{N}\sum_{j=1}^{N}\xi^{(f)}_j\xi^{(q)}_j + \mathrm{O}(N^{-3/2}). \tag{12}$$

Remark. If a_N and b are proper random variables then the asymptotic relation $a_N = N^\gamma b + \mathrm{o}(N^\gamma)$ means that $\lim_{N\to\infty} N^{-\gamma}a_N \stackrel{d}{=} b$. The asymptotic relation $a_N = \mathrm{O}(N^\gamma)$ means that $\lim_{N\to\infty} N^{-\gamma}a_N$ is a proper random variable.

Note that in (11) and (12) we have written explicitly the terms of the order $\mathrm{O}(N^{-1/2})$ in the expression for $x_p(n)$ and the terms of the order $\mathrm{O}(1)$ and $\mathrm{O}(N^{-1})$ in the expression for $x_q(n)$ (there are no terms $\sim aN^{-1/2}$ in the latter case). This choice is motivated by two reasons. First, the evolution of the explicitly written terms is closed. Neither $O(N^{-1})$ terms in $x_p(n)$, nor $O(N^{-3/2})$ terms in $x_q(n)$ interfere in the evolution of the explicitly written terms. Second, only the explicitly written terms make non-vanishing contributions to $N\Phi_N(x_1, \ldots, x_M)$ in the limit $N \to \infty$. Moreover, the contributions to $N\Phi_N(x_1, \ldots, x_M)$ from the terms of the order $O(N^{-1/2})$ in $x_p(n)$ and from the terms of the order $O(N^{-1})$ in $x_q(n)$ have the same order of magnitude.

To find the fixed point of the recurrent relations (11) and (12) note that the large N asymptotic expansion of $x_p(n)$ are given by

$$x_p(n) = \frac{1}{\sqrt{N}}\delta_p(n) + o(N^{-1/2}) \tag{13}$$

for $p \neq q$, and

$$x_q(n) = x_* + \frac{1}{N}\delta_q(n) + o(N^{-1}) \tag{14}$$

where $\delta_q(n)$ are random variables independent of N distributions.

Substituting the expansions (13) and (14) in (11) and (12) one obtains the following recurrent relations for $\delta_p(n)$:

$$\delta_p(n+1) = x_*\mathcal{N}_{q,p}(0;1) + \frac{\beta}{\cosh^2(\beta x_*)}\delta_p(n) \tag{15}$$

if $p \neq q$, and

$$\delta_q(n+1) = \frac{\beta}{\cosh^2(\beta x_*)}\left[\delta_q(n) + \sum_{f(\neq q)}^{M} \delta_f(n)\mathcal{N}_{f,q}(0;1)\right]. \tag{16}$$

The coordinates of the fixed point of these recurrent relation are given by

$$\delta_p^* = \frac{x_*}{1-\beta/\cosh^2(\beta x_*)}\mathcal{N}_{q,p}(0;1)$$

for $p \neq q$, and

$$\delta_q^* = \frac{1}{[1-\beta/\cosh^2(\beta x_*)]^2}\frac{\beta x_*}{\cosh^2(\beta x_*)}\sum_{f(\neq q)}^{M}\mathcal{N}_{f,q}^2(0;1)$$

where

$$\mathcal{N}_{q,p}(0;1) \stackrel{d}{=} \lim_{N\to\infty}\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\xi_j^{(p)}\xi_j^{(q)}$$

are random variables with the standard normal distribution. Therefore, the initial condition $\boldsymbol{x}_q^*$ is in the basin of attraction of the fixed point $\boldsymbol{x}_{k,N}^* = (x_1^{(q)}, x_2^{(q)}, \ldots, x_M^{(q)})$ with the coordinates

$$x_p^{(q)} = \frac{1}{\sqrt{N}}\frac{x_*}{1-\beta/\cosh^2(\beta x_*)}\mathcal{N}_{q,p}(0;1) + o\left(\frac{1}{\sqrt{N}}\right)$$

for $p \neq q$, and

$$x_q^{(q)} = x_* + N^{-1}\frac{1}{[1-\beta/\cosh^2(\beta x_*)]^2}\frac{\beta x_*}{\cosh^2(\beta x_*)}\sum_{f(\neq q)}^{M}\mathcal{N}_{f,q}^2(0;1) + o(N^{-1}).$$

Taking into account the above expressions for $x_p^{(q)}$, $p = 1, 2, \ldots, M$; one obtains

$$\Phi(\boldsymbol{x}_{q,N}^*) = \frac{1}{2}\beta x_*^2 - \log\cosh(\beta x_*) - \frac{\beta x_*^2 N^{-1}}{2(1-\beta+\beta x_*^2)}\sum_{f(\neq q)}^{M}\mathcal{N}_{f,q}^2(0;1) + o(N^{-1})$$

as $N \to \infty$. Application of the Laplace method to the integral (5) (in $2M$ points of minima $\pm \boldsymbol{x}^*_{q,N}$, $q = 1, 2, \ldots, M$) yields

$$Z_N e^{\frac{1}{2}\beta M} = 2^{N+1}(1-\beta+\beta x_*^2)^{-M/2} \exp[-N(\tfrac{1}{2}\beta x_*^2 - \log\cosh(\beta x_*))] \times \sum_{q=1}^{M} \exp\left[\frac{\beta x_*^2}{2(1-\beta+\beta x_*^2)} \sum_{f(\neq q)}^{M} \mathcal{N}^2_{f,q}(0;1)\right](1+o(1)). \tag{17}$$

The random variables

$$\nu_q \equiv \sum_{f(\neq q)}^{M} \mathcal{N}^2_{f,q}(0;1) \qquad q = 1, 2, \ldots, M$$

have χ^2 distribution with $M-1$ degrees of freedom, that is, their common distribution density is given by

$$f_n(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2} & \text{if } x > 0 \\ 0 & \text{if } x \leqslant 0 \end{cases} \tag{18}$$

where $n = M - 1$, see figure 2. The random variables ν_q are dependent since $\mathcal{N}_{p,q}(0;1) \equiv \mathcal{N}_{q,p}(0;1)$. Therefore, (apparently) no simple expression for the distribution of the partition function for arbitrary M can be found and the usefulness of equation (17) deteriorates very rapidly as M increases. However, in the limit $M \to \infty$ the distribution of Z_N simplifies. Indeed, one has

$$\lim_{M\to\infty} \mathcal{N}_q^{(M)} \equiv \lim_{M\to\infty} \frac{1}{\sqrt{M-1}} \sum_{f(\neq q)}^{M} [\mathcal{N}^2_{f,q}(0;1) - 1] \stackrel{d}{=} \mathcal{N}_q(0;1) \tag{19}$$

as $M \to \infty$, where the random variables $\mathcal{N}_q(0;1)$ have the standard normal distribution (see figure 2 for illustration). Note that the random variables $\{\mathcal{N}_q^{(M)}\}_{q=1}^{M}$ are not independent but exchangeable, for instance,

$$\boldsymbol{E}(\mathcal{N}_k^{(M)}\mathcal{N}_j^{(M)}) = \frac{1}{M-1} \sum_{f(\neq k)} \sum_{l(\neq j)} \boldsymbol{E}[(\mathcal{N}^2_{f,k}(0;1) - 1)(\mathcal{N}^2_{l,j}(0;1) - 1)] = \frac{2}{M-1}$$

for $k \neq j$. According to equation (17)

$$Z_N e^{\frac{1}{2}\beta M} = 2^{N+1}(1-\beta+\beta x_*^2)^{-M/2} \exp[-N(\tfrac{1}{2}\beta x_*^2 - \log\cosh(\beta x_*))] \times \exp\left[\frac{\beta x_*^2 (M-1)}{2(1-\beta+\beta x_*^2)}\right] \mathcal{R}_M \tag{20}$$

where

$$\mathcal{R}_M \equiv \sum_{q=1}^{M} \exp\left[\frac{\beta x_*^2}{2(1-\beta+\beta x_*^2)} \sqrt{M-1}\,\mathcal{N}_q^{(M)}\right]$$

is the 'random' part of the partition function. To find the distribution of $\mathcal{R}_M$ for large M we use the double inequality

$$\max_{k=1,2,\ldots,M} \mathcal{N}_k^{(M)} \leqslant \frac{1}{\gamma\sqrt{M-1}} \log\left[\sum_{k=1}^{M} \exp\left(\gamma\sqrt{M-1}\,\mathcal{N}_k^{(M)}\right)\right] \leqslant \max_{k=1,2,\ldots,M} \mathcal{N}_k^{(M)} + \frac{\log M}{\gamma\sqrt{M-1}}. \tag{21}$$

For large M the joint distribution density of the random variables $\{\mathcal{N}_q^{(M)}\}_{q=1}^M$, is approximated by its quadratic (Gaussian) part

$$g(x_1, x_2, \ldots, x_M) = \frac{\sqrt{\det \hat{C}}}{(2\pi)^{M/2}} \exp\left[-\frac{M-1}{2(M-3)} \sum_{j=1}^{M} x_j^2 + \frac{1}{3(M-3)} \left(\sum_{j=1}^{M} x_j\right)^2\right]$$

where $\hat{C}$ is the covariance matrix of the random variables $\{\mathcal{N}_q^{(M)}\}_{q=1}^M$, and $\det \hat{C} = \frac{1}{3}(\frac{M-1}{M-3})^{M-1}$. Using the identity (4) one obtains

$$\begin{aligned}\Pr\left[\max_{k=1,\ldots,M} \mathcal{N}_k^{(M)} \leqslant x\right] &= \Pr[\mathcal{N}_1^{(M)} \leqslant x, \ldots, \mathcal{N}_M^{(M)} \leqslant x] \\ &\approx \frac{\sqrt{\det \hat{C}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}y\, \mathrm{e}^{-\frac{1}{2}y^2} \left\{ \int_{-\infty}^{x} \frac{\mathrm{d}\tau}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{M-1}{M-3}\tau^2 + \frac{\sqrt{2}\tau y}{\sqrt{3(M-3)}}\right]\right\}^M \\ &= \frac{1}{\sqrt{6\pi}}\sqrt{\frac{M-3}{M-1}} \int_{-\infty}^{\infty} \mathrm{d}y\, \exp\left[-\frac{y^2}{6}\left(1 - \frac{2}{M-1}\right)\right] I^M(x, y)\end{aligned}$$

where

$$I(x, y) \equiv \int_{-\infty}^{b(x,y)} \frac{\mathrm{d}\tau}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\tau^2}$$

and

$$b(x, y) = x\sqrt{1 + \frac{2}{M-3}} - y\sqrt{\frac{2}{3(M-1)}}.$$

Using a well known result of probability theory one has for $y = 0$

$$I^n\left(\sqrt{2\log n} - \frac{\log(4\pi \log n)}{2\sqrt{2\log n}} + \frac{\chi}{\sqrt{2\log n}}, 0\right) \to \exp(-\mathrm{e}^{-\chi})$$

as $n \to \infty$. The same is obviously true for any y. Hence

$$\Pr\left[\max_{k=1,\ldots,M} \mathcal{N}_k^{(M)} \leqslant \sqrt{2\log M} - \frac{\log(4\pi \log M)}{2\sqrt{2\log M}} + \frac{\chi}{\sqrt{2\log M}}\right] \to \exp(-\mathrm{e}^{-\chi})$$

as $M \to \infty$. Thus,

$$\begin{aligned}\mathcal{R}_M = \exp\Bigg\{\frac{\beta x_*^2}{2(1-\beta+\beta x_*^2)}&\left[\sqrt{2M\log M} - \frac{1}{2}\sqrt{\frac{M}{2\log M}}\log(4\pi \log M) + \sqrt{\frac{M}{2\log M}}r\right] \\ &+ \mathrm{o}\left(\sqrt{\frac{M}{2\log M}}\right)\Bigg\} \qquad (22)\end{aligned}$$

where r is a random variable with the iterated exponential distribution

$$\Pr[r \leqslant x] = \exp(-\mathrm{e}^{-x}).$$

The main features of the iterated exponential distribution can be seen in figure 1.

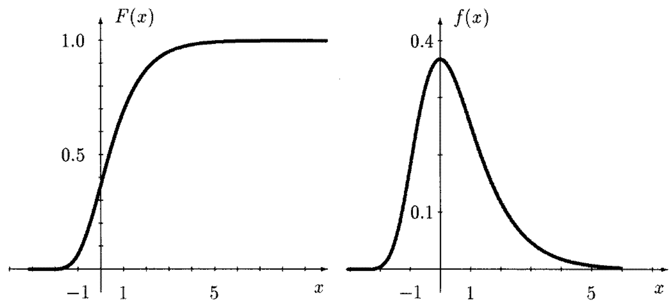

Figure 1. The iterated exponential distribution $F(x) = \exp(-\mathrm{e}^{-x})$ and its density $f(x) = \exp(-\mathrm{e}^{-x} - x)$.

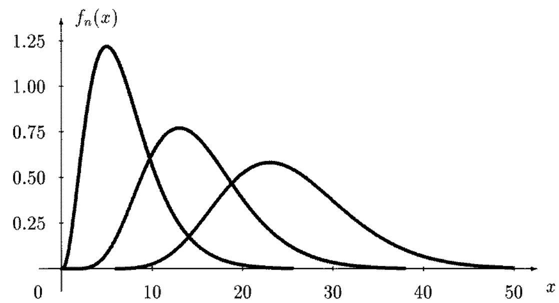

Figure 2. The χ^2 distribution densities with 7, 15 and 25 degrees of freedom.

2.2. The high-temperature region

For $\beta \in (0; \beta_c)$, where $\beta_c = 1$, the function $\Phi(\boldsymbol{x})$ given by equation (7) has a unique minimum at the point $\boldsymbol{x} = (0, \ldots, 0)$. Indeed, the stationary points of $\Phi(\boldsymbol{x})$ are solutions of

$$x_p = \frac{1}{2^M} \sum_{\sigma_1, \ldots, \sigma_M = \pm 1} \tanh\left[\beta\left(x_p + \sigma_p \sum_{f(\neq p)}^{M} \sigma_f x_f\right)\right] \qquad p = 1, 2, \ldots, M.$$

This system has an obvious solution $\boldsymbol{x} = (0, \ldots, 0)$. Using the elementary inequality

$$\tfrac{1}{2}[\tanh\beta(x-y) + \tanh\beta(x+y)] \leqslant \tanh\beta x \qquad \text{for } x \geqslant 0$$

one has

$$\frac{1}{2^M} \sum_{\sigma_1, \ldots, \sigma_M = \pm 1} \tanh\left[\beta\left(\sigma_p \sum_{f(\neq p)}^{M} \sigma_f x_f + x_p\right)\right] \leqslant \tanh\beta x_p$$

for $x_p \geqslant 0$. It follows then that $(0, \ldots, 0)$ is the only minimum of $\Phi(x_1, \ldots, x_M)$ for $\beta \in (0; 1)$ since $\tanh\beta x < x$ for $\beta \in (0; 1)$ and $x > 0$.

Since the first partial derivatives of $\Phi_N(\boldsymbol{x})$ converge to the first partial derivatives of $\Phi(\boldsymbol{x})$ all stationary points of the function $\Phi_N(\boldsymbol{x})$ are in an arbitrarily small vicinity of the point $(0, \ldots, 0)$ when N is large enough. The function $\Phi(\boldsymbol{x})$ is convex and its matrix of second derivatives is negatively defined in a neighbourhood of the point $(0, \ldots, 0)$. The same is true for the functions $\Phi_N(\boldsymbol{x})$, when N is large enough, since their second partial derivatives converge to the second partial derivatives of $\Phi(\boldsymbol{x})$. Therefore the function $\Phi_N(\boldsymbol{x})$ has a unique minimum at the point $(0, \ldots, 0)$ when N is large enough.

Application of the Laplace method for the evaluation of the integral in equation (5) yields a deterministic leading-order term in the asymptotic expansion of the partition function. It is given by

$$Z_N = 2^N(1-\beta)^{-M/2}\mathrm{e}^{-M\beta/2}(1+\mathrm{O}(N^{-1})) \tag{23}$$

for $\beta \in (0; 1)$. Obviously, the partition function (3) still has a non-degenerate distribution in the high-temperature region (with the exception of the case $M = 1$), however, all 'randomness' is now hidden in the $\mathrm{O}(N^{-1})$ term.

Note that non-trivial disordered systems, like the Sherrington–Kirkpatrick model and the Hopfield model with macroscopic number of patterns, do not share this property. Namely, the main asymptotics of the corresponding partition functions have non-degenerate distributions even in the high temperature regions [1, 11]. The partition function of the random energy model has, however, deterministic leading-order term in the high-temperature region, see [5].

2.3. The critical point

At the critical point $\beta = 1$ the function $\Phi(\boldsymbol{x})$ still has the unique minimum at the point $(0, \ldots, 0)$. However, all its second partial derivatives vanish at that point and the matrix of the second derivatives of the function $\Phi_N(\boldsymbol{x})$ may not be negatively defined at the the point $(0, \ldots, 0)$ depending on the realization of randomness. Therefore the exact locations of the stationary points of the function $\Phi_N(\boldsymbol{x})$ and even their number are random for $\beta = 1$. Keeping only the relevant terms one has

$$\Phi_N(x_1; \ldots; x_M) \approx -\frac{1}{N}\sum_{p<f}^{M}\sum_{j=1}^{N}\xi_j^{(p)}\xi_j^{(f)}x_p x_f + \frac{1}{12}\sum_{p=1}^{M}x_p^4 + \frac{1}{2}\sum_{p<f}^{M}x_p^2 x_f^2.$$

It is apparently impossible to find explicit expressions for the stationary points. However, it is clear that the coordinates of the stationary points of the function $\Phi_N(\boldsymbol{x})$ scale with N as $x_p = N^{-1/4}\chi_p$, where χ_p are random variables with non-degenerate distributions. Application of the Laplace method shows that the leading-order term in the large-N asymptotic expansion of the partition function at the critical point has a non-degenerate distribution.

To complete our investigation of the partition function corresponding to the Hamiltonian (2) we now consider the case $\varepsilon > 0$. In this case one can still easily obtain an integral representation for the partition function analogous to equation (5). In fact one simply has to replace the function $\Phi_N(\boldsymbol{x})$ in (5) by

$$\mathcal{F}_N(x_1, \ldots, x_M) = \frac{1}{2}\beta\sum_{p=1}^{M}x_p^2 - \frac{1}{N}\sum_{i=1}^{N}\log\cosh\left[\beta\left(\sum_{p(\neq q)}^{M}x_p\xi_i^{(p)} + (x_q+\varepsilon)\xi_i^{(q)}\right)\right].$$

For any $\varepsilon > 0$ only the minimum $\boldsymbol{x}_{q,N}^*$ close to the point $\boldsymbol{x}_q^* = \{\delta_{l,q}x_*(\varepsilon)\}_{l=1}^M$, where $x_*(\varepsilon)$ is the maximal solution of $x = \tanh[\beta(x+\varepsilon)]$, makes contribution to the main asymptotics of the partition function. Repeating the arguments which we used to obtain equation (20) one arrives at

$$\begin{aligned} Z_N\mathrm{e}^{\frac{1}{2}\beta M} &= 2^N(1-\beta+\beta x_*^2(\varepsilon))^{-M/2}\exp[-N(\tfrac{1}{2}\beta x_*^2(\varepsilon) - \log\cosh(\beta x_*(\varepsilon)))] \\ &\quad\times\exp\left[\frac{\beta x_*^2(\varepsilon)}{2(1-\beta+\beta x_*^2(\varepsilon))}\sum_{f(\neq q)}^{M}\mathcal{N}_{f,q}^2(0;1) + \mathrm{o}(1)\right]. \end{aligned} \tag{24}$$

Thus the distribution of the main asymptotic of the partition function in the case $\varepsilon \neq 0$ is governed by the χ^2 distribution with $M - 1$ degrees of freedom (main features of the χ^2 distributions are illustrated in figure 2). For large M the (properly rescaled) χ^2 distributions with $M-1$ degrees of freedom approach the normal distribution (see equation (19)), therefore, in the case $\varepsilon \neq 0$ the large M asymptotics of the random part of the partition function has distribution governed by the normal distribution.

As one could have expected the limit $\varepsilon \downarrow 0$ in (24) does not yield (20).

3. Concluding remarks

In the present paper we found the probability distribution of the main asymptotics of the partition function of the Hopfield model with a finite number of patterns (for all temperatures except the critical point). The main computational steps which allow one to obtain explicit expressions are application of the Laplace method for the evaluation of the multiple integral in equationuation (5) and solution of the equations for stationary points (9). It is of course interesting to know if one can obtain similar results for Hopfield models with infinite (in the thermodynamic limit) number of patterns. Recent results for Hopfield models with the number of patterns $M = \mathrm{o}(N)$ [7, 12], suggest that the formal application of the Laplace method still yields the correct expression for the leading-order term in the large-N asymptotic expansion of the corresponding partition functions. The explicit expressions for the location of the stationary points can still be found by the method of the present paper if $M = \mathrm{o}(N)$. Therefore, all the results of the present paper can be rederived in the case $M = \mathrm{o}(N)$. It is clear, however, that they coincide with the $M \to \infty$ asymptotics of the results found in the present paper. That is, the formal substitution $M \to N^\gamma$ in (22) and (23) yields the expressions valid in the case $M = N^\gamma$, $\gamma \in (0; 1)$.

When $M = \alpha N$ the Laplace method needs some modification since the major contribution to the main asymptotics of the partition function does not necessarily come from the global minima of the function $\Phi_N(\boldsymbol{x})$. The curvature of this function at a stationary point becomes important when $M = \alpha N$. What becomes a really difficult problem is the location of the stationary points. The method of the present paper obviously cannot be applied for this problem in the case $M = \alpha N$. However, it can be used to find an expansion in powers of α for the corresponding stationary points (which is, possibly, only an asymptotic series).

Acknowledgments

The author would like to thank one of the referees for several useful comments on the presentation of the paper and for spotting errors in some equations.

References

[1] Aizenman M., Lebowitz J L and Ruelle D 1987 Some rigorous results on the Sherrington–Kirkpatrick spin-glass model *Commun. Math. Phys.* **112** 3–20

[2] Amaro de Matos J M G, Patrick A E and Zagrebnov V A 1992 Random infinite-volume Gibbs states for the Curie–Weiss random field Ising model *J. Stat. Phys.* **66** 139–64

[3] Bovier A 1994 Self-averaging in a class of generalized Hopfield models *J. Phys. A: Math. Gen.* **27** 7069–77

[4] Bovier A, Gayrard V and Picco P 1994 Gibbs states of the Hopfield model in the regime of perfect memory *Prob. Theor. Rel. Fields* **100** 329–63

[5] Galves A, Martinez S and Picco P 1989 Fluctuations in Derrida's random energy and generalized random energy models *J. Stat. Phys.* **54** 515–29

[6] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci., USA* **79** 2554–8
[7] Koch H 1993 A free energy bound for the Hopfield model *J. Phys. A: Math. Gen.* **26** L353–5
[8] Pastur L A and Figotin A L 1978 On the theory of disordered spin systems *Theor. Math. Phys.* **35** 403–14
[9] Pastur L A and Shcherbina M V 1991 Absence of self-averaging of the order parameter in the Sherrington–Kirkpatrick model *J. Stat. Phys.* **62** 1–19
[10] Pastur L, Shcherbina M and Tirozzi B 1994 The replica-symmetric solution without replica trick for the Hopfield model *J. Stat. Phys.* **74** 1161–83
[11] Scacciatelli E and Tirozzi B 1992 Fluctuations of the free energy in the Hopfield model *J. Stat. Phys.* **67** 981–1008
[12] Shcherbina M and Tirozzi B 1992 The free energy for a class of Hopfield models *J. Stat. Phys.* **72** 113–25