THE FLUCTUATIONS OF THE OVERLAP IN THE HOPFIELD MODEL WITH FINITELY MANY PATTERNS AT THE CRITICAL TEMPERATURE

BARBARA GENTZ AND MATTHIAS LÖWE

ABSTRACT. We investigate the limiting fluctuations of the order parameter in the Hopfield model of spin glasses and neural networks with finitely many patterns at the critical temperature $1/\beta_c = 1$. At the critical temperature, the measure-valued random variables given by the distribution of the appropriately scaled order parameter under the Gibbs measure converge weakly towards a random measure which is non-Gaussian in the sense that it is not given by a Dirac measure concentrated in a Gaussian distribution. Moreover, we analyze the limiting behaviour of the fluctuations as the temperature $1/\beta_N$ converges to one and obtain a threshold for the speed of this convergence to one: If β_N converges to $\beta_c = 1$ at speed $o(1/\sqrt{N})$, then the non-Gaussian behaviour of the fluctuations persists, whereas for β_N converging to one at a speed which is slower than $\mathcal{O}(1/\sqrt{N})$, the order parameter obeys a weak central limit theorem.

1. Introduction and Description of the Model

The Hopfield model, together with the Sherrington–Kirkpatrick model of spin glasses, is one of the most popular models of disordered spin systems. The origin of the interest in it is at least twofold. On the one hand, it can be regarded as a very simple model of the brain, hence as a neural network. This point of view has been taken in the original paper by Hopfield [15] for instance, as well as in the papers [21], [22], [17], [19], [20], and many others. On the other hand, the Hopfield model may also be regarded as a spin glass with the additional feature that the complexity of the disorder can be adjusted by an additional parameter, the number of patterns. This has been the motivation for a variety of papers including [23], where the Hopfield model has been first introduced, and [24], [5], [6], [7], [1], [11], [12], [2], [3], and [26] to list some of the papers in this area (of course, this list is far from being complete). For an overview of recent results on the Hopfield model and related models, see [8] and, in particular, [4] therein.

To be more specific, let us define the Hopfield model. First of all we fix a number $M \in \mathbb{N}$ which will denote the number of so-called patterns, which in contrast to the most general set-up of the model will not depend on the number $N \in \mathbb{N}$ of spins or "neurons". By

$$H_{N,\xi}(\sigma) = -\frac{1}{2N} \sum_{\mu=1}^{M} \sum_{i,j=1}^{N} \sigma_i \sigma_j \xi_i^\mu \xi_j^\mu, \qquad \sigma \in \{-1,+1\}^N, \tag{1.1}$$

Date: May 9, 1998.

1991 *Mathematics Subject Classification.* 60F05, 60K35 (primary), 82C32 (secondary).

Key words and phrases. Hopfield model, spin glasses, neural networks, limit theorems, central limit theorem, Gaussian and non-Gaussian fluctuations, critical temperature.

we will denote the (random) Hamiltonian of the Hopfield model, which we will regard as a function of the spin configuration $\sigma \in \{-1, +1\}^N$. The variable $\xi_i^\mu \in \{-1, +1\}$ stands for the ith component of the μth pattern. We will assume that the ξ_i^μ are i.i.d. unbiased random variables, i. e., that

$$\mathbb{P}(\xi_i^\mu = +1) = \mathbb{P}(\xi_i^\mu = -1) = \frac{1}{2}.$$

Expectations with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}$. For convenience, we denote by ξ the $(N \times M)$-matrix consisting of the $(\xi_i^\mu)_{i,\mu}$, while $\xi_i = (\xi_i^1, \dots, \xi_i^M)$ and $\xi^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)$, respectively, stand for the ith row and the μth column of this matrix, respectively.

The spin variables' a priori distribution P is assumed to be unbiased as well, i. e.,

$$\mathrm{P}(\sigma_i = +1) = \mathrm{P}(\sigma_i = -1) = \frac{1}{2}$$

for all $i \in \mathbb{N}$. Expectations with respect to the spin variables' a priori distribution P will be denoted by E. We shall assume throughout this note that the family $\{\, \xi_i^\mu : i \in \{1, \dots, N\}, \mu \in \{1, \dots, M\} \,\}$ is independent of the family of the spin variables $\{\, \sigma_i : i \in \{1, \dots, N\} \,\}$.

The Hopfield model at temperature $1/\beta \in (0, \infty)$ may now either be identified with the Hamiltonian (1.1) or, equivalently, with the Gibbs measure

$$\varrho_{N,\beta,\xi}(\sigma) = 2^{-N} \exp\{-\beta H_{N,\xi}(\sigma)\}/Z_{N,\beta}(\xi), \qquad \sigma \in \{-1, +1\}^N,$$

where the so-called partition function

$$Z_{N,\beta}(\xi) = \frac{1}{2^N} \sum_{\sigma \in \{-1,+1\}^N} \exp\{-\beta H_{N,\xi}(\sigma)\}$$

is the appropriate normalization to make $\varrho_{N,\beta,\xi}$ a probability measure.

Note that the Hamiltonian (1.1) may be rewritten in the following convenient form as a quadratic functional of the overlap m_N:

$$H_{N,\xi}(\sigma) = -\frac{N}{2} \|m_N(\sigma)\|_2^2,$$

where

$$m_N(\sigma) = (m_N^\mu(\sigma))_{\mu=1,\dots,M} \qquad \text{with} \qquad m_N^\mu(\sigma) = \sum_{i=1}^N \xi_i^\mu \sigma_i.$$

Here and below, $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^M$. The μth component of the overlap compares the spin configuration to the μth pattern ξ^μ. Therefore, the overlap m_N is an important quantity for the investigation of the Hopfield model, a so-called order parameter. Its distribution under $\varrho_{N,\beta,\xi}$ has been of major interest in recent studies of the model. In [5], the asymptotic behaviour of the overlap under the Gibbs measure $\varrho_{N,\beta,\xi}$ was first investigated. The authors established a law of large numbers which holds for $\mathbb{P}$-almost all realizations of the random patterns ξ. They show that for $\mathbb{P}$-almost all ξ, the distribution of the overlap m_N under the Gibbs measure with external magnetic field of strength $h \neq 0$ in the direction of the first unit vector e_1 of the canonical basis in $\mathbb{R}^M$ converges weakly towards the Dirac measure $\delta_{\pm z(\beta) e_1}$ concentrated in $\pm z(\beta) e_1$ as first the system size $N \to \infty$ and then the strength $h \to 0\pm$. Here $z(\beta)$ denotes the largest root $z \in [0, 1)$ of the Curie–Weiss equation

$$z = \tanh(\beta z).$$