MODERATE DEVIATIONS FOR THE OVERLAP PARAMETER IN THE HOPFIELD MODEL

PETER EICHELSBACHER AND MATTHIAS LÖWE

ABSTRACT. We derive moderate deviation principles for the overlap parameter in the Hopfield model of spin glasses and neural networks. If the inverse temperature β is different from the critical inverse temperature $\beta_c = 1$ and the number of patterns $M(N)$ satisfies $M(N)/N \to 0$, the overlap parameter multiplied by N^γ, $1/2 < \gamma < 1$, obeys a moderate deviation principle with speed $N^{1-2\gamma}$ and a quadratic rate function (i.e. the Gaussian limit for $\gamma = 1/2$ remains visible on the moderate deviation scale). At the critical temperature we need to multiply the overlap parameter by N^γ, $1/4 < \gamma < 1$. If then $M(N)$ satisfies $(M(N)^6 \log N \wedge M(N)^2 N^{4\gamma} \log N)/N \to 0$, the rescaled overlap parameter obeys a moderate deviation principle with speed $N^{1-4\gamma}$ and a rate function that is basically a fourth power. The random term occurring in the Central Limit theorem for the overlap at $\beta_c = 1$ is no longer present on a moderate deviation scale. If the scaling is even closer to $N^{1/4}$, e.g. if we multiply the overlap parameter by $N^{1/4} \log\log N$ the moderate deviation principle breaks down. The case of variable temperature converging to one is also considered. If β_N converges to β_c fast enough, i.e. faster than $\mathcal{O}(N^{-2\gamma})$, the non-Gaussian rate function persists, whereas for β_N converging to one slower than $\mathcal{O}(N^{-2\gamma})$, the moderate deviations principle is given by the Gaussian rate. At the borderline the moderate deviation rate function is the one at criticality plus an additional Gaussian term.

1. INTRODUCTION

In 1976 and 1977, respectively Sherrington and Kirkpatrick on the one hand and Pastur and Figotin on the other introduced and discussed disordered versions of the Curie–Weiss model of ferromagnets (see [29], [27], [28]). These models — nowadays known under the names Sherrington–Kirkpatrick (SK) model and Hopfield model — are two of the most popular models of spin glasses. While the SK model still has resisted attempts to rigorously solve it for large regimes of its parameters the Hopfield model nowadays is much better understood, at least when it exhibits paramagnetic or magnetic behaviour. Since its re-interpretation by Hopfield [18] the latter also has played a prominent role in the theory of neural networks. This versatility of the Hopfield model—namely that it can be regarded as a very simple model of the brain on one hand, and as a disordered spin system on the other hand—has led research on the Hopfield model into two different directions.

1991 *Mathematics Subject Classification.* 60F10 (primary), 60K35, 82B44, 82D30 (secondary).

Key words and phrases. moderate deviations, large deviations, Hopfield model, neural networks, spin glasses, critical temperature, random disorder.

Research supported by the Volkswagen-Stiftung (RiP-program at Oberwolfach, Germany).

The neural network point of view has been taken in the original paper by Hopfield [18] for instance, as well as in the papers [24], [25], [19], [22], [23], and many others. The spin glass aspects of the model on the other hand were studied in the seminal paper [27], as well as in [26], [5], [1], [14], [30], [2], [20], [17], [6], [31], [32] among others. Of course, giving a complete list of all (important) papers in this area is almost impossible. A good overview of many aspects of the model is provided e.g. in [7], in particular, [4] therein.

Let us now define the Hopfield model. First of all $N \in \mathbb{N}$ will denote the number of spins or "neurons" while $M(N)$ stands for the number of the so-called patterns. We will be interested in how $M = M(N)$ may depend on N. We will write

$$\alpha := \alpha(N) := \frac{M(N)}{N}$$

and throughout this paper we will be in the situation where $\alpha(N) \searrow 0$. The random function

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

denotes the Hamiltonian of the Hopfield model, which is a function of the spin configuration $\sigma \in \{-1,+1\}^N$.

The strength of the pair interaction $\sum_{\mu=1}^{M(N)} \xi_i^\mu \xi_j^\mu$ is random as the variables $\xi_i^\mu \in \{-1,+1\}$ with ξ_i^μ (denoting the i'th component of the μ'th pattern) are random. More precisely we shall assume that the ξ_i^μ are i.i.d. random variables such that for given N, the family of random variables $\{ \xi_i^\mu : i \in \{1,\dots,N\}, \mu \in \{1,\dots,M(N)\} \}$ is independent with

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

for all i and μ. Expectations with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}$. Whenever convenient, we shall write ξ for the $(N \times M(N))$-matrix consisting of the $(\xi_i^\mu)_{i,\mu}$. $\xi_i = (\xi_i^1, \dots, \xi_i^{M(N)})$ and $\xi^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)$, respectively, denote the i'th row and the μ'th column of this matrix, respectively.

The spin variables are a priori assumed to be independent and independent of the family $\{ \xi_i^\mu : i \in \{1,\dots,N\}, \mu \in \{1,\dots,M(N)\} \}$ with an unbiased distribution P, i.e.

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

for all $i \in \mathbb{N}$.

The Hopfield model at temperature $1/\beta \in (0,\infty)$ may now be identified with the Gibbs measure with respect to the Hamiltonian (1.1), i. e.,

$$\varrho_{N,\beta}(\sigma) = 2^{-N} \exp\{-\beta H_N(\sigma)\}/Z_{N,\beta}, \qquad \sigma \in \{-1,+1\}^N, \tag{1.2}$$

where the so-called partition function

$$Z_{N,\beta} = \frac{1}{2^N} \sum_{\sigma \in \{-1,+1\}^N} \exp\{-\beta H_N(\sigma)\} \tag{1.3}$$

is the normalization which makes $\varrho_{N,\beta}$ a probability measure.

Note that the Hamiltonian (1.1) may be rewritten as a quadratic functional of the so-called overlap m_N:

$$H_N(\sigma) = -\frac{N}{2}\|m_N(\sigma)\|_2^2, \tag{1.4}$$

where

$$m_N(\sigma) = (m_N^\mu(\sigma))_{\mu=1,\dots,M(N)} \qquad \text{with} \qquad m_N^\mu(\sigma) = \frac{1}{N}\sum_{i=1}^N \xi_i^\mu \sigma_i. \tag{1.5}$$

Here and below, $\|\cdot\|_2 = \|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^{M(N)}$. The μ'th component $m_N^\mu(\sigma)$ of the overlap $m_N(\sigma)$ compares the spin configuration to the μ'th pattern ξ^μ in such a way that a large absolute value of $m_N^\mu(\sigma)$ means that the spin configuration σ largely agrees with ξ^μ (or its negative). These configurations are of low energy according to (1.4). Therefore, the overlap is an important quantity for the investigation of the Hopfield model, a so-called order parameter. Its distribution under $\varrho_{N,\beta}$ has been of major interest in the study of the model and also will be central in this paper.

In [5], a law of large numbers for the distribution of the overlap under the Gibbs measure $\varrho_{N,\beta}$ is established. This law was shown to hold for $\mathbb{P}$-almost all realizations of the random patterns ξ, whenever $M(N)/N \to 0$. More precisely, the authors in [5] prove that for $M(N)/N \to 0$ and 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 in $\mathbb{R}^{M(N)}$ (i.e. the Gibbs measure with Hamiltonian $H_N(\sigma)+h\sum_i \xi_i^1\sigma_i$) converges weakly toward the Dirac measure $\delta_{z^\pm(\beta)e_1}$ concentrated in $z^\pm(\beta)e_1$ as first $N\to\infty$ and then $h \to 0\pm$. Here $z^+(\beta)$ denotes the largest root $z \in [0,1)$ of the Curie–Weiss equation

$$z = \tanh(\beta z). \tag{1.6}$$

Note that $z^+(\beta) = 0$ for $\beta \le \beta_c = 1$. Therefore δ_0 is the unique limiting measure for the distribution of $m_N(\sigma)$ in the high-temperature regime $\beta \le \beta_c = 1$. On the other hand $z^+(\beta) > 0$ for $\beta > \beta_c$. Hence in the low temperature regime the limit point will not be unique.

The corresponding large deviation principle (LDP for short) was established in [1]. For almost all ξ the sequence $(m_N)_{N\in\mathbb{N}}$ under the Gibbs measure $\varrho_{N,\beta}$ obeys a principle of large deviations with speed N and *deterministic* rate function I. The rate function I in general is not convex.

Here we say that a sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$, on some topological space X obeys a large deviation principle with speed a_n and good rate function $I(\cdot): X \to \mathbb{R}_0^+$ if

- I is lower semi-continuous and has compact level sets $N_L := \{x \in X : I(x) \le L\}$, for every $L \in [0,\infty)$.
- For every open set $G \subseteq X$ it holds

$$\liminf_{n\to\infty} \frac{1}{a_n}\log \mu_n(G) \ge -\inf_{x\in G} I(x). \tag{1.7}$$

- For every closed set $A \subseteq X$ it holds

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \mu_n(A) \leq - \inf_{x\in A} I(x). \tag{1.8}$$

Similarly, we will say that a sequence of random variables $(Y_n)_{n\in\mathbb{N}}$ with topological state space X obeys a large deviation principle with speed a_n and good rate function $I(\cdot) : X \to \mathbb{R}_0^+$ if the sequence of their distributions does. Formally a moderate deviation principle is nothing but an LDP. However, we will speak about a moderate deviation principle (MDP) for a sequence of random variables, whenever the scaling of the corresponding random variables is between that of an ordinary Law of Large Numbers and that of a Central Limit Theorem.
On the scale of fluctuations, when analyzing the distribution of $\sqrt{N}(m_N - z(\beta)e_1)$, the disorder becomes visible. Indeed, for $M(N)/N \to 0$ and $(\beta, h) \neq (1,0)$, the overlap under $\varrho_{N,\beta}$ satisfies $\mathbb{P}$-almost surely a central limit theorem with the covariance matrix which could be expected from the analogy with the Curie–Weiss model and a centering which differs in the cases $\beta > 1$ or $h \neq 0$ from the naively expected one by a ξ-dependent adjustment, see [14], [13], [15] and [2].
As in the Curie–Weiss model at the critical temperature $(\beta, h) = (1,0)$ the fluctuations are non–Gaussian. Even more the limiting distribution has a random component. In [17], [16], and [32] it was shown that in the Hopfield model with a number of patterns $M(N)$ such that $M(N)^{13}\log(N)^\gamma/N \to 0$ for some $\gamma > 0$ the following convergence result holds. The distribution of the overlap—scaled by the factor $N^{1/4}$—converges weakly (with respect to $\mathbb{P}$) to a limiting random measure $\overline{Q}_{M(N)}$. This limiting random measure $\overline{Q}_{M(N)}$ is defined by density with respect to the $M(N)$-dimensional Lebesgue measure which is proportional to

$$\exp\left(-\frac{1}{12}\sum_{\mu=1}^{M(N)} x_\mu^4 - \frac{1}{2} \sum_{1\leq\mu<\nu\leq M(N)} x_\mu^2 x_\nu^2 + \sum_{1\leq\mu<\nu\leq M(N)} \eta_{\mu,\nu} x_\mu x_\nu \right). \tag{1.9}$$

Here η is an $M(N)(M(N)-1)/2$-dimensional Gaussian random variable with mean zero and the covariance matrix being the identity matrix. This shows that at the critical temperature $\beta_c = 1$, the fluctuations of the overlap depend strongly on the random disorder as even the distribution of the limiting fluctuations is random.
In this paper we are aiming to prove a moderate deviation principle for the overlap parameter m_N under $\varrho_{N,\beta}$. It is kind of folklore in large deviation theory that typically the rate function in a large deviations regime will depend on the distribution of the underlying random variables, while an MDP inherits properties of both the central limit behaviour as well as the large deviation principle: the speed of convergence to zero of the probability of an atypical event usually is exponential while the rate function does not depend on the fine structure of the underlying distribution and is determined by the large deviation behaviour of the limiting distribution in the Central Limit Theorem.
However, we will see in the present article that this folklore is not always true. Indeed, we will find that even if a Central Limit Theorem and an LDP hold true,

this does not necessarily imply an MDP, and even if such an MDP holds the rate function may be slightly surprising.

Motivated by our results in [11] and the central limit theorems quoted above, we expect the moderate deviations behaviour to be completely different for the two cases $(\beta, h) \neq (1,0)$ and $(\beta, h) = (1,0)$. In fact, in the cases we study, this will turn out to be true. We will only consider the case of zero external field, i.e. $h \equiv 0$, and expect the case of $h \neq 0$ to be very similar to the high temperature case $\beta < 1$ and $h = 0$.
To be able to state our results denote by $\Pi_k^{M(N)}$ the projection operator

$$\Pi_k^{M(N)} : \mathbb{R}^{M(N)} \to \mathbb{R}^k.$$

For convenience we will also write (in a slight abuse of notation) Π_k instead of $\Pi_k^{M(N)}$ ignoring that the dimension of $\mathbb{R}^{M(N)}$ might be growing.
Then for the moderate deviations of m_N under $\varrho_{N,\beta}$ we have the following results:

Theorem 1.1. *[Moderate deviations, $\beta \neq 1$]*
Assume that $M(N)/N \to 0$.
For $\beta < 1$, any fixed k, any $0 < \gamma < 1/2$ and $\mathbb{P}$-almost all sequences ξ the sequence of measures $\varrho_{N,\beta} \circ (\Pi_k(N^\gamma m_N))^{-1}$ obeys a moderate deviations principle with speed $N^{1-2\gamma}$ and rate function

$$I(x) := I_k(x) := \frac{1-\beta}{2} \sum_{i=1}^{k} x_i^2. \tag{1.10}$$

For $\beta > 1$, any fixed k, any $0 < \gamma < 1/2$ and $\mathbb{P}$-almost all sequences ξ there exists an ε_0 and a random centering $\overline{X}_{N,\beta}^{l,\varepsilon} := \overline{X}_{N,\beta}^{l,\varepsilon}(\xi)$ such that for all $\varepsilon < \varepsilon_0$ the sequence of measures

$$\varrho_{N,\beta}\left(\Pi_k\left[N^\gamma m_N - \overline{X}_{N,\beta}^{l,\varepsilon}\right] \in \bullet \middle| \, \|m_N - z^+(\beta) e_l\| < \varepsilon\right) \tag{1.11}$$

satisfies a moderate deviations principle with speed $N^{1-2\gamma}$ and rate function

$$I(x) := I_k(x) := \frac{1-\beta(1-z^+(\beta)^2)}{2(1-z^+(\beta)^2)} \sum_{i=1}^{k} x_i^2. \tag{1.12}$$

Here e_l is the l'th unit vector, $l = 1, \ldots, M(N)$, and $z^+(\beta)$ is the largest solution of the Curie–Weiss equation (1.6). The same holds true, if we replace $z^+(\beta)$ by $z^-(\beta)$.

For $\beta = 1$ and a scaling not too close to the corresponding CLT we are also able to prove an MDP. The rate function this time is given by the deterministic terms in the exponent of the corresponding density function in the CLT. More precisely, the result reads as follows.

Theorem 1.2. *[Moderate deviations, $\beta = 1$]*
Assume that $M(N)^6(\log N)/N \to 0$ as well as $M(N)^2(\log N)/N^{1-4\gamma} \to 0$ for the γ introduced below. For $\beta = 1$, any fixed k, any $0 < \gamma < 1/4$, and $\mathbb{P}$-almost all sequences ξ the sequence of measures $\varrho_{N,\beta} \circ (\Pi_k(N^\gamma m_N))^{-1}$ obeys a moderate deviations principle with speed $N^{1-4\gamma}$ and rate function

$$I(x) := I_k(x) := \frac{1}{12}\sum_{i=1}^k x_i^4 + \frac{1}{2}\sum_{1\le\mu<\nu\le k} x_\mu^2 x_\nu^2. \tag{1.13}$$

Remark 1.3. Note that the conditions on the growth rate of $M(N)$ in Theorems 1.1 is the same as the condition in Corollary 1.2 in [2] which is the best known growth rates on $M(N)$ for the corresponding Central Limit Theorem. At $\beta = 1$ the situation is more involved. If the MDP scale is not too close to the corresponding CLT scale our condition on the growth rate of $M(N)$ is better than the one in [16], but worse than that given by Talagrand for a similar result in [32, Theorem 1.1]. Of course it would be nice to obtain an MDP at $\beta = 1$ at least for his growth condition on $M(N)$. However, an adaption of his methods is at least difficult. Note that in his Theorem 1.1 in [32] he approximates probabilities (for m_N being in a subset) with an error of $1/N^2$ which is far from sufficient to control moderate deviation probabilities. On the other hand decreasing this error to an exponential scale automatically gives worse conditions on $M(N)$ (see [32, Lemma 3.6]).
In fact, the situation is even worse. Even if we could adapt Talagrand's methods, we would not be able to obtain his condition on $M(N)$. As will turn out soon, the MDP will break down for a scaling very close to the CLT regime for *any* choice of $M(N) > 1$, i.e. even if $M(N)$ is constant. This makes it very plausible that also for other scalings the admissible $M(N)$ will depend on the scale as they do in our Theorem 1.2 (we do not claim that this dependence is optimal but some sort of dependency will very likely exist).

As just mentioned, if our scaling gets too close to the CLT at $\beta = 1$, the MDP breaks down, for any growth rate of $M(N)$. This is stated in the following

Theorem 1.4. *Let $\beta = 1$ and let b_N increase to infinity such that*

$$b_N \ll \sqrt[4]{\log\log N}$$

in the sense of 1.5 below. Let $M(N)$ be weakly increasing (i.e. possibly eventually constant in N) and $M(1) > 1$. Then for any fixed k and $\mathbb{P}$-almost all sequences ξ the sequence of measures $\varrho_{N,\beta} \circ \left(\Pi_k(\frac{N^{1/4}}{b_N} m_N)\right)^{-1}$ does not obey an MDP.

We also have a result for β_N depending on the system size N:

Notation 1.5. For two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ with $a_n, b_n \geq 0$ let us write $a_n \ll b_n$ if $a_n \leq b_n$ and $a_n/b_n \to 0$ and, similarly, we shall write $a_n \gg b_n$ if $b_n \ll a_n$.

Theorem 1.6. *[Moderate deviations for size-dependent temperatures] Assume that* $M(N)^6(\log N)/N \to 0$ *as well as* $M(N)^2(\log N)/N^{1-4\gamma} \to 0$ *for the* γ *introduced below. Moreover let* $0 < \beta_N < \infty$ *depend on* N *in such a way that* $\beta_N \to 1$ *monotonically as* $N \to \infty$. *Then the following assertions hold for any fixed* k, *any* $0 < \gamma < 1/4$ *and* $\mathbb{P}$*-almost all sequences* ξ:

- *If* $|\beta_N - 1| \ll 1/N^{2\gamma}$, *the sequence of measures* $\varrho_{N,\beta} \circ (\Pi_k(N^\gamma m_N))^{-1}$ *obeys a moderate deviations principle with speed* $N^{1-4\gamma}$ *and rate function*

$$I(x) := I_k(x) := \frac{1}{12}\sum_{i=1}^k x_i^4 + \frac{1}{2}\sum_{1\le\mu<\nu\le k} x_\mu^2 x_\nu^2.$$

- *If* $|\beta_N - 1| \gg 1/N^{2\gamma}$, *the sequence of measures* $\varrho_{N,\beta} \circ (\Pi_k(N^\gamma m_N))^{-1}$ *obeys a moderate deviations principle with speed* $|1-\beta_N|N^{1-2\gamma}$ *and rate function*

$$I(x) := I_k(x) := \frac{1}{2}\sum_{i=1}^k x_i^2.$$

- *If* $|\beta_N - 1| = \delta\big(1/N^{2\gamma}\big)$, *for some* $\delta > 0$, *the sequence of measures* $\varrho_{N,\beta} \circ (\Pi_k(N^\gamma m_N))^{-1}$ *obeys a moderate deviations principle with speed* $N^{1-4\gamma}$ *and rate function*

$$I(x) := I_k(x) := \frac{1}{12}\sum_{i=1}^k x_i^4 + \frac{1}{2}\sum_{1\le\mu<\nu\le k} x_\mu^2 x_\nu^2 + \frac{\delta}{2}\sum_{i=1}^k x_i^2.$$

This paper contains two more sections. In Section 2 we will prove Theorem 1.1. This proof will mainly follow ideas presented in [2]. Section 3 contains the proof of Theorems 1.2, 1.4 and 1.6. It is basically inspired by the ideas used in [16].

2. The case of $\beta \neq 1$ — Proof of Theorem 1.1

This sections contains the proof of Theorem 1.1. We will restrict ourselves to low temperature case $\beta > 1$, the proof in the other case being just a rerun of the arguments we will present below. As already mentioned these arguments are inspired by the work in [2].

Let us denote

$$\mathcal{R}_{N,\beta}^{l,\varepsilon}(A) := \varrho_{N,\beta}\left(m_N \in A \,\middle|\, \|m_N - z^+(\beta)e_l\| < \varepsilon\right), \qquad A \in \mathcal{B}(\mathbb{R}^{M(N)}). \tag{2.1}$$

For $\varepsilon > 0$ and $l \in \{1,\dots,M(N)\}$ denote by B_ε^l the ball of radius ε centered in $z^+(\beta)e_l$. Our goal is to establish an MDP for appropriately normalized random variables $(X_N)_{N\in\mathbb{N}}$ drawn according to $\mathcal{R}_{N,\beta}^{l,\varepsilon}$. This MDP will be shown by considering the corresponding *Laplace transform* of the measures of interest and by applying the Theorem of Gärtner and Ellis (see for example [9, Theorem 2.3.6]). So for X_N

sampled from $\mathcal{R}^{l,\varepsilon}_{N,\beta}$ denote by $\overline{X}^{l,\varepsilon}_{N,\beta}$ its expectation. Then the object of interest is the *Laplace transform*

$$L^{l,\varepsilon}_{N,\beta}(t) := \int \exp\left(N^{1-\gamma}\langle t, x - \overline{X}^{l,\varepsilon}_{N,\beta}\rangle\right) d\mathcal{R}^{l,\varepsilon}_{N,\beta}(x), \qquad t \in \mathbb{R}^{M(N)}, \tag{2.2}$$

where $\langle\cdot,\cdot\rangle$ stands for the inner product in $\mathbb{R}^{M(N)}$. We prove the following theorem.

Theorem 2.1. *Assume that* $\lim_{N\to\infty}\alpha(N) \to 0$ *and assume that* $\beta > 1$. *Denote by*

$$\lambda(\beta) := [1-\beta(1-z^+(\beta)^2)].$$

Let

$$r \geq \max\left\{N^{-\gamma/2}, 2\sqrt{\frac{\alpha(N)}{\lambda(\beta)}}\right\}. \tag{2.3}$$

Then we have for $\mathbb{P}$*-almost all sequences* ξ *and for* t *with* $\|t\| < \infty$ *that the Laplace transform* $L^{l,r}_{N,\beta}(t)$ *satisfies*

$$\lim_{N\to\infty}\frac{1}{N^{1-2\gamma}}\log L^{l,r}_{N,\beta}(t) = \frac{\|t\|^2}{2}\frac{1-z^+(\beta)^2}{1-\beta(1-z^+(\beta)^2)}. \tag{2.4}$$

Corollary 2.2. *Consider a sequence of random variables* $(X_N)_{N\in\mathbb{N}}$ *distributed according to* $\mathcal{R}^{l,\varepsilon}_{N,\beta}$. *Then for any fixed* r *satisfying* (2.3), *any fixed* $k \in \mathbb{N}$ *and all* t *with* $\|t\| < \infty$ *we have for* $\mathbb{P}$*-almost all sequences* ξ *that the Laplace transform*

$$L^{l,\varepsilon}_{N,\beta,k}(t) := \int \exp\left(\Pi_k\left(N^{1-\gamma}\langle t, z - \overline{X}^{l,\varepsilon}_{N,\beta}\rangle\right)\right) d\mathcal{R}^{l,\varepsilon}_{N,\beta}(z) \tag{2.5}$$

satisfies

$$\lim_{N\to\infty}\frac{1}{N^{1-2\gamma}}\log L^{l,r}_{N,\beta,k}(t) = \frac{\|\Pi_k(t)\|^2}{2}\frac{1-z^+(\beta)^2}{1-\beta(1-z^+(\beta)^2)}. \tag{2.6}$$

Proof of Theorem 1.1. The Theorem now follows from Corollary 2.2 by an application of the Gärtner-Ellis theorem [9, Theorem 2.3.6]. The rate function is obtained as the Legendre–Fenchel transform of $\frac{\|\Pi_k(t)\|^2}{2}\frac{1-z^+(\beta)^2}{1-\beta(1-z^+(\beta)^2)}$, which is indeed computed to be $I(\cdot)$ as given in (1.12). □

To prove Theorem 2.1, in the first place, instead of considering our original sequence of measures $\varrho_{N,\beta}\circ(m_N)^{-1}$ we will consider a smoothed version of these measures, i.e. we apply the so called *Hubbard–Stratonovich transform*. This idea was successfully used in the context of the analysis of the Curie–Weiss model in e.g. [12] and already introduced to the analysis of the Hopfield model in [27]. It consists of convoluting our measures of interest with $\mathcal{N}(0, \frac{1}{\beta N}\mathcal{I}_{M(N)})$, i.e. with an $M(N)$-dimensional Gaussian measure with mean 0 and covariances matrix $\frac{1}{\beta N}$ times the identity matrix.

Lemma 2.3. *Denote*

$$\mathcal{Q}_{N,\beta} = \left(\varrho_{N,\beta}\circ(m_N)^{-1}\right) * \mathcal{N}\left(0, \frac{1}{\beta N}\mathcal{I}_{M(N)}\right).$$

Then $\mathcal{Q}_{N,\beta}$ is absolutely continuous with respect to $M(N)$-dimensional Lebesgue measure and has density

$$\frac{1}{\Xi} \exp\left(-\beta N \Phi_{N,\beta,\xi}(z)\right). \tag{2.7}$$

Here $\Phi_{N,\beta,\xi}(z) =: \Phi_N(z)$ is given by

$$\Phi_{N,\beta,\xi}(z) := \frac{1}{2}\|z\|^2 - \frac{1}{\beta N} \sum_{i=1}^{N} \log\cosh(\beta\langle \xi_i, z\rangle), \quad z \in \mathbb{R}^{M(N)}, \tag{2.8}$$

with

$$\Xi := \int_{\mathbb{R}^{M(N)}} \exp\left(-\beta N \Phi_{N,\beta,\xi}(z)\right) dz$$

is the proper normalization to make $\mathcal{Q}_{N,\beta}$ a probability measure.

Proof. The proof of this and similar statements can be found at many places, e.g. in [5, Lemma 2.2]. □

For the measures $\mathcal{Q}_{N,\beta}$ we will now introduce their conditional versions. We define

$$\mathcal{Q}^{l,\varepsilon}_{N,\beta}(A) := \mathcal{Q}_{N,\beta}\left(A \,\middle|\, B^l_\varepsilon\right), \qquad A \in \mathcal{B}(\mathbb{R}^{M(N)}). \tag{2.9}$$

We will start by deriving an MDP for (appropriately normalized) random variables $(Y_N)_{N\in\mathbb{N}}$ drawn according to $\mathcal{Q}^{l,\varepsilon}_{N,\beta}$. This MDP will be shown by considering the corresponding Laplace transform of the measures of interest. So for Y_N sampled from $\mathcal{Q}^{l,\varepsilon}_{N,\beta}$ denote by $\overline{Y}^{l,\varepsilon}_{N,\beta}$ its expectation. Denote by

$$\mathcal{L}^{l,\varepsilon}_{N,\beta}(t) := \int \exp\left(N^{1-\gamma}\langle t, y - \overline{Y}^{l,\varepsilon}_{N,\beta}\rangle\right) d\mathcal{Q}^{l,\varepsilon}_{N,\beta}(y), \qquad t \in \mathbb{R}^{M(N)}, \tag{2.10}$$

the corresponding Laplace transform.
We are now ready to prove our first (and most important) Lemma on the sequence $(Y_N)_{N\in\mathbb{N}}$. Let us remind that $0 < \gamma < 1/2$ and $\alpha(N) := M(N)/N$.

Lemma 2.4 (cf.[2, Proposition 2.2])**.** *Assume that $\beta \neq 1$. Let $\lambda(\beta)$ be as in Theorem 2.1 and asumme that*

$$r(N) \geq \max\left\{N^{-\gamma/2}, 2\sqrt{\frac{\alpha(N)}{\lambda(\beta)}}\right\} \tag{2.11}$$

is decreasing in N such that $r(N) \searrow 0$ as $N \to \infty$. Then for all but a finite number of N and all $t \in \mathbb{R}^{M(N)}$ the Laplace transform $\mathcal{L}^{l,r(N)}_{N,\beta}(t)$ satisfies: for $\mathbb{P}$-almost all ξ there is a sequence $\delta_N \searrow 0$ as $N \to \infty$ such that

$$\frac{N^{1-2\gamma}\|t\|^2 \left(1 - 3r(N)e^{-M(N)}\right)}{2\beta\left(\lambda(\beta) + \delta_N\right)} \leq \log \mathcal{L}^{l,r(N)}_{N,\beta}(t) \leq \frac{N^{1-2\gamma}\|t\|^2 \left(1 + 2r(N)e^{-M(N)}\right)}{2\beta\left(\lambda(\beta) - \delta_N\right)}. \tag{2.12}$$

Remark 2.5. Under the same conditions as in Theorem 2.1 the estimates (2.12) will give us

$$\lim_{N\to\infty} \frac{1}{N^{1-2\gamma}} \log \mathcal{L}_{N,\beta}^{l,r(N)}(t) = \frac{\|t\|^2}{2\beta\lambda(\beta)} = \frac{\|t\|^2}{2\beta[1-\beta(1-z^+(\beta)^2)]}, \tag{2.13}$$

saying that the MDP for the projection of the sequence $(Y_N)_{N\in\mathbb{N}}$ drawn form $\mathcal{Q}_{N,\beta}^{l,\varepsilon}$ holds true with a rate function that is the Legendre-Fenchel transform of $\frac{\|\Pi_k(t)\|^2}{2\beta[1-\beta(1-z^+(\beta)^2)]}$.

The proof of Lemma 2.4 follows the ideas of Bovier and Gayrard in [2, Lemma 2.2 and Lemma 2.1] and is based on the following Lemma, which is a version of the well known *Brascamp–Lieb inequality* and its reverse, see [8] and [10]. Let T be a $M(N)\times M(N)$ real symmetric matrix. T is said to be non-negative, and we write $T\geq 0$, if $\langle x, Tx\rangle \geq 0$ for any $x\in\mathbb{R}^{M(N)}$. For any function $V:\mathbb{R}^{M(N)}\to\mathbb{R}$, we will denote by $\nabla^2 V(x)$ its Hessian matrix in x.

Lemma 2.6. *Let $r(N), \lambda(\beta)$ and δ_N be as in Lemma 2.4. Let $V:\mathbb{R}^{M(N)}\to\mathbb{R}$ be a positive function such that for all $x\in B_{r(N)}^l$, $l\in\{1,\dots,M(N)\}$,*

$$0 < (\lambda(\beta)-\delta_N)\mathcal{I} \leq \nabla^2 V(x) \leq (\lambda(\beta)+\delta_N)\mathcal{I}, \tag{2.14}$$

where $\mathcal{I}$ denotes the $M(N)\times M(N)$-identity matrix.
Then, denoting by $\mathbb{E}_V$ the expectation with respect to the probability measure on $\big(B_{r(N)}^l, \mathcal{B}(B_{r(n)}^l)\big)$, $l\in\{1,\dots,M(N)\}$,

$$\mathbb{P}_V(dx) := \frac{e^{-NV(x)}\, 1_{\{x\in B_{r(N)}^l\}}}{\int_{B_{r(N)}^l} e^{-NV(x)} d^{M(N)}x}\, d^{M(N)}x$$

one gets for any $t\in\mathbb{R}^{M(N)}$,

$$\frac{1}{2\lambda(\beta)+\delta_N} N^{1-2\gamma}\|t\|^2 - R_N \leq \log \mathbb{E}_V e^{N^{1-\gamma}\langle t, X-\mathbb{E}_V(X)\rangle} \leq \frac{1}{2\lambda(\beta)-\delta_N} N^{1-2\gamma}\|t\|^2 + R_N. \tag{2.15}$$

Here

$$R_N = \frac{2\, r(N)\, e^{-M(N)}}{\lambda(\beta)-\delta_N} N^{1-2\gamma}\|t\|^2.$$

Condition (2.14) says that the function V is assumed to be strictly convex and moreover to have a Hessian which is uniformly close to a multiple of the identity. The Brascamp-Lieb inequalities together with a reverse introduced in [10] now yield asymptotically coinciding upper and lower bounds on the Laplace transform. The smoothed versions $\mathcal{Q}_{N,\beta}$ of $\varrho_{N,\beta}\circ(m_n)^{-1}$ have the density (2.7). Now Lemma 2.1 in [2] tells us that condition (2.14) is fulfilled for $\beta\,\Phi_{N,\beta,\xi}$.

Proof of Lemma 2.6. The proof is a rerun of the arguments leading to Lemma 2.2 in [2], the only difference being that our exponent is $s(X) = N^{1-\gamma}\langle t, X-\mathbb{E}_V(X)\rangle$

while there it is $\overline{s}(X) = N^{1/2}\langle t, X - \mathbb{E}_V(X)\rangle$. Both functions are linear in X and hence obey

$$\nabla^2\big(V(X) + s(X)\big) = \nabla^2\big(V(X) + \overline{s}(X)\big) = \nabla^2 V(X).$$

This shows that the proof of Lemma 2.2 in [2] works. □

Proof of Lemma 2.4. Lemma 2.4 is an immediate consequence of Lemma 2.6. Indeed, [2, Lemma 2.1] shows that the function $\beta\,\Phi_{N,\beta,\xi}$ satisfies (2.14) on balls of radius $r(N)$ centered in 0 as long as $r(N)$ tends to zero with N tending to infinity. The lower bound can already be found in [3] and [4], the prove of the upper bound is given in detail in [2, Lemma 2.1], using a careful study of the structure of the minima of the random function $\Phi_{N,\beta,\xi}$. (2.12) follows from Lemma 2.6. □

Lemma 2.4 would (almost) enable us to prove an MDP for the sequence $\big(Y_N - \overline{Y}^{l,\varepsilon}_{N,\beta}\big)_{N\in\mathbb{N}}$, see Remark 2.5. Unfortunately, this is not we claimed. We want to derive an MDP for a sequence $\big(X_N - \overline{X}^{l,\varepsilon}_{N,\beta}\big)_{N\in\mathbb{N}}$, where X_N is drawn from $\mathcal{R}^{l,\varepsilon}_{N,\beta}$ (defined as in (2.1)) and $\overline{X}^{l,\varepsilon}_{N,\beta}$ is its expectation. Along the lines of the proof of Proposition 2.1 in [2] this is done in three steps. First, we show that our result for the logarithmic Laplace transform for radii converging to zero can be transferred to fixed radii.

Lemma 2.7. *Assume that $\beta > 1$. For any $r(N)$ and any fixed r satisfying* (2.11) *and all $t \in \mathbb{R}^{M(N)}$ we have for all but a finite number of indices N and for $\mathbb{P}$-almost all sequences ξ*

$$\mathcal{L}^{l,r(N)}_{N,\beta}(t)(1 - e^{-cM(N)}) \le \mathcal{L}^{l,r}_{N,\beta}(t) \le e^{-c(M(N)\wedge N^{1-\gamma})} + \mathcal{L}^{l,r(N)}_{N,\beta}(t)(1 + e^{-cM(N)}) \quad (2.16)$$

for some constant $c > 0$, which may depend on $\beta > 1$.

Proof. Already Bovier and Gayrard prove in [2, Proposition 2.1 (ii)] that (2.16) is true with $e^{-cM(N)}$ instead of $e^{-c(M(N)\wedge N^{1-\gamma})}$. The first inequality in (2.16) is proved exactly as on page 170 in [2], where their estimate (3.41) is shown. An easy adaption of the calculations in (3.42) and (3.43) in [2] shows that the upper bound for $\tilde{T}_2$, defined in [2, (3.39)], changes to

$$\tilde{T}_2 \le \exp\Big(-c_1\beta M(N) + N^{1-\gamma}\|t\|\tilde{\varrho} + \frac{\|t\|^2}{\beta c_2(z^+(\beta))^2}\Big) \times$$
$$\exp\left(-\gamma\beta N c_2(z^+(\beta))^2\Big(r(N) - \frac{\|t\|}{\beta N^{1-\gamma}c_2(z^+(\beta))^2}\Big)^2\right)\left(\frac{1}{(z^+(\beta))^2(1-\gamma)}\right)^{M(N)/2},$$

for any $0 < \gamma < 1$, where c_1, c_2 are constants depending only on β and $\tilde{\varrho}$ is chosen to be $\tilde{\varrho} := c_3\sqrt{\alpha(N)}$. Since our radii are potentially larger than $\Big(N^{-\gamma/2} \wedge \sqrt{\frac{\alpha(N)}{\lambda(\beta)}}\Big)$, and since $r(N) > \tilde{\varrho}$ the corresponding Gaussian integrals concentrate faster and yield the bound

$$0 \le \tilde{T}_2 \le \exp\big(-c(M(N) \wedge N^{1-\gamma})\big)$$

for N large enough and for some constant $c > 0$, which gives the second inequality in (2.16). □

Next we replace the centering $\overline{Y}^{l,\varepsilon}_{N,\beta}$ by the centering $\overline{X}^{l,\varepsilon}_{N,\beta}$.

Lemma 2.8. *For any r satisfying (2.3) and all $t \in \mathbb{R}^{M(N)}$ we have for all but a finite number of indices N and for $\mathbb{P}$-almost all sequences ξ*

$$\left|\langle \overline{Y}^{l,r}_{N,\beta} - \overline{X}^{l,r}_{N,\beta}, t\rangle\right| \le \|t\|\, e^{-c\left(M(N)\wedge N^{1-\gamma}\right)}$$

for some constant $c = c(\beta) > 0$. Therefore

$$\left|\log \mathcal{L}^{l,r}_{N,\beta}(t) - \log \int \exp\left(N^{1-\gamma}\langle t, y - \overline{X}^{l,r}_{N,\beta}\rangle\right) d\mathcal{Q}^{l,r}_{N,\beta}(y)\right| \le N^{1-\gamma}\|t\| e^{-c\left(M(N)\wedge N^{1-\gamma}\right)}. \tag{2.17}$$

Proof. This is nothing but Proposition 2.1 (iii) in [2], where (2.17) is true with $e^{-cM(N)}$ instead of our $e^{-c(M(N)\wedge N^{1-\gamma})}$. Modifications as in the proof of Lemma 2.7 give the result. □

Eventually we are able to treat the Laplace transform of the sequence $(X_N)_{N\in\mathbb{N}}$

Lemma 2.9. *For any fixed r satisfying (2.3) and all $t \in \mathbb{R}^{M(N)}$ we have for all but a finite number of indices N and for $\mathbb{P}$-almost all sequences ξ that the Laplace transform $L^{l,\varepsilon}_{N,\beta}$ defined in (2.2) behaves in the following way*

$$L^{l,r}_{N,\beta}(t)(1-\varepsilon_N) \le e^{-N^{1-2\gamma}\|t\|^2/2\beta} \int \exp\left(N^{1-\gamma}\langle t, y - \overline{X}^{l,r}_{N,\beta}\rangle\right) d\mathcal{Q}^{l,r}_{N,\beta}(y) \tag{2.18}$$

and

$$\begin{aligned} e^{-N^{1-2\gamma}\|t\|^2/2\beta} \int \exp\left(N^{1-\gamma}\langle t, y - \overline{X}^{l,r}_{N,\beta}\rangle\right) d\mathcal{Q}^{l,r}_{N,\beta}(y) \\ \le e^{-c(M(N)\wedge N^{1-\gamma})} + L^{l,r}_{N,\beta}(t)(1+\varepsilon_N) \end{aligned} \tag{2.19}$$

for some sequence $\varepsilon_N > 0$ converging to zero as N goes to infinity.

Proof. A careful analysis of the proof of Proposition 2.1 (i) in [2] shows that a proper choice of their δ_N^2 ([2, (3.28)]) gives our bound. We shortly point out the changes of their proof, see pages 166–170. Their estimate (3.19) ends with

$$\frac{e^{-N^{1-\gamma}\langle t, \overline{X}^{l,r}_{N,\beta}\rangle}}{Z_{N,\beta}} E\left(e^{(1/2)\beta N\|m_N(\sigma)+t/(N^\gamma)\|^2}\right) \times \mathcal{N}^{M(N)}_{\beta N}\left(B_r\left(z^+(\beta)e_l - \left(m_N(\sigma) + \frac{t}{\beta N^\gamma}\right)\right)\right).$$

Here $\mathcal{N}^{M(N)}_{\beta N}$ denotes the Gaussian measure on $(\mathbb{R}^{M(N)}, \mathcal{B}(\mathbb{R}^{M(N)}))$ with density

$$\left(\frac{\beta N}{2\pi}\right)^{M(N)} \exp\left(-\frac{1}{2}\beta N\|z\|^2\right)$$

and $B_r(x)$ denotes a ball in $\mathbb{R}^{M(N)}$ around x with radius r. With the same techniques as on page 167 and 168 in [2] one gets

$$e^{-N^{1-2\gamma}\|t\|^2/2\beta}\int \exp\left(N^{1-\gamma}\langle t, y-\overline{X}^{l,r}_{N,\beta}\rangle\right) d\mathcal{Q}^{l,r}_{N,\beta}(y) \leq \tag{2.20}$$
$$e^{-(1/2)\beta N\delta_N^2+4N^{1-\gamma}\|t\|+(1/2)\beta M(N)}$$
$$+\quad L^{l,r}_{N,\beta}(t)\left(1+e^{-c(\beta)\beta M(N)}+2e^{-(1/2)\beta N^{(2\gamma)}\delta_N^2+(1/2)\beta M(N)}\right).$$

Distinguishing the two cases $M(N) < N^{1-\gamma}$ and $M(N) > N^{1-\gamma}$, we choose

$$\delta_N^2 := \begin{cases} (\beta+8\|t\|+2c(\beta))\frac{1}{\beta N^\gamma}, & \text{if } M(N) < N^{1-\gamma},\\ (\beta+8\|t\|+2c(\beta))\frac{M(N)}{\beta N^{2\gamma}} & \text{if } M(N) > N^{1-\gamma}.\end{cases}$$

Inserting this choice of δ_N^2 into (2.20) yields our result. □

Proof of Theorem 2.1. The proof of Theorem 2.1 now follows immediately. By Lemma 2.7 the difference between the logarithms of the Laplace transforms for different r goes to zero fast enough. Hence we obtain (2.13) for any fixed r satisfying (2.3). The difference arising from the different centering goes to zero on the right scale by Lemma 2.8. Finally, the original Laplace transform $L^{l,r}_{N,\beta}$ is related to $\int \exp\left(N^{1-\gamma}\langle t, y-\overline{X}^{l,\varepsilon}_{N,\beta}\rangle\right) d\mathcal{Q}^{l,r}_{N,\beta}(y)$ by the estimates in Lemma 2.9, for any fixed r satisfying (2.3). Hence together with Lemma 2.4 we obtain

$$\lim_{N\to\infty}\frac{1}{N^{1-2\gamma}}\log L^{l,r}_{N,\beta}(t) = -\frac{\|t\|^2}{2\beta}+\frac{\|t\|^2}{2\beta[1-\beta(1-z^+(\beta)^2)]} = \frac{\|t\|^2}{2}\frac{1-z^+(\beta)^2}{1-\beta(1-z^+(\beta)^2)},$$

and therefore Theorem 2.1 is proven. □

3. The critical case — Proof of Theorems 1.2, 1.4, and 1.6

In this section we will turn to the case of critical temperature. Other than in the previous section our proofs now are inspired by [16]. Similar to the previous section we begin with convoluting our measure of interest with a Gaussian measure.

Lemma 3.1. *Let* $0<\beta<\infty$ *and* $a_N>0$ *for every* $N\in\mathbb{N}$. *Then the convolution*

$$\mathcal{Q}_{N,\beta,a_N} = \left(\varrho_{N,\beta}\circ(\sqrt{a_N}m_N)^{-1}\right) * \mathcal{N}\left(0,\frac{a_N}{\beta N}\mathcal{I}_{M(N)}\right)$$

is the random measure on $\mathbb{R}^{M(N)}$ *which is given by the (random) density*

$$f_{N,\beta,a_N}(x) := \frac{1}{\Xi}\exp\left(-\beta N\Phi_{N,\beta,\xi}(x/\sqrt{a_N})\right), \qquad x\in\mathbb{R}^{M(N)}, \tag{3.1}$$

with

$$\Xi = \int_{\mathbb{R}^{M(N)}}\exp\left(-\beta N\Phi_{N,\beta,\xi}(x/\sqrt{a_N})\right)dx$$

and $\Phi_{N,\beta,\xi}$ *defined as in* (2.8).

The proof is very similar to the one (not) given in Section 2.
We will apply Lemma 3.1 for $\beta = 1$ (or $\beta_N \to 1$ for the proof of Theorem 1.6) and for $a_N = N^{2\gamma}$, $0 < \gamma < 1/4$. Other than in Section 2 the MDP behaviour of $\mathcal{Q}_{N,1,N^{2\gamma}}$ will be completely identical with the MDP behaviour of $\varrho_{N,\beta} \circ (N^\gamma m_N)^{-1}$. This will be proven in the following Lemma.

Lemma 3.2. *Under the conditions and with the notation of Lemma 3.1, let $(X_N)_{N\in\mathbb{N}}$ be a sequence of random variables distributed according to $\mathcal{Q}_{N,1,N^{2\gamma}}$. If for some fixed $k \in \mathbb{N}$ the sequence $(\Pi_k(X_N))_{N\in\mathbb{N}}$ satisfies an MDP with speed $N^{1-4\gamma}$ and some rate function H for $\mathbb{P}$-almost all sequences ξ, then so does $(\Pi_k(N^\gamma m_N))_{N\in\mathbb{N}}$ (where $N^\gamma m_N$ is distributed according to $\varrho_{N,1} \circ (N^\gamma m_N)^{-1}$) and the speed and rate functions agree.*

Proof. Note that $X_N = N^\gamma m_N + Y_N$ in distribution, where $(Y_N)_{N\in\mathbb{N}}$ is a sequence of (independent) Gaussian random variables on $\mathbb{R}^{M(N)}$ with mean zero and covariance matrix $\frac{1}{N^{1-2\gamma}}\mathcal{I}_{M(N)}$ and independent of the sequence $(m_N)_{N\in\mathbb{N}}$. Hence for any fixed $k \in \mathbb{N}$ and any $\varepsilon > 0$

$$\mathrm{Prob}\big(|\Pi_k(X_N - N^\gamma m_N)| \geq \varepsilon\big) = \mathrm{Prob}\big(|\Pi_k(Y_N)| \geq \varepsilon\big) \leq C \exp\big(-N^{1-2\gamma}\varepsilon^2\big).$$

Here Prob denotes a probability measure that makes Y_N and m_N independent, $N \in \mathbb{N}$. This implies

$$\limsup_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \mathrm{Prob}\left(|\Pi_k(X_N - N^\gamma m_N)| > \varepsilon\right) = -\infty,$$

showing that the two sequences $(\Pi_k(X_N))_{N\in\mathbb{N}}$ and $(\Pi_k(N^\gamma m_N))_{N\in\mathbb{N}}$ are exponentially equivalent for $\mathbb{P}$-almost all ξ and therefore have the same moderate deviation behaviour (see [9, Theorem 4.2.13]). ☐

Next we will Taylor-expand the function $\Phi_N(x)$, defined in (2.8) around zero.

Lemma 3.3. *$\Phi_N(x)$ has the following Taylor expansion around zero: There exists a $\theta \in (0,1)$ such that for $x \in \mathbb{R}^{M(N)}$*

$$\Phi_N(x) = \Phi_{N,\beta,\xi}(x) = \frac{1}{2}\|x\|_2^2 - \frac{1}{\beta N}\sum_{i=1}^N \left[\frac{\beta^2}{2}\langle x, \xi_i\rangle^2 - \frac{\beta^4}{12}\langle x, \xi_i\rangle^4\right] + R_{N,\beta}(x,\xi), \quad (3.2)$$

with

$$R_{N,\beta}(x,\xi) = -\frac{1}{N\beta}\sum_{i=1}^N \frac{\beta^5}{15} h(\langle \theta x, \xi_i\rangle)\langle x, \xi_i\rangle^5, \quad (3.3)$$

where

$$h(t) = \frac{\tanh(t)}{\cosh^4(t)}[2 - \sinh^2(t)], \qquad t \in \mathbb{R}.$$

In particular we find for $N\,\Phi_N(x/N^\gamma)$ and for $\beta = 1$ that:

$$
\begin{aligned}
&-N\,\Phi_N(x/N^\gamma)\\
&\quad = N^{1-4\gamma}\Bigg(-\frac{1}{12}\|x\|_4^4 - \frac{1}{4}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}^2 x_{\mu_2}^2 + \frac{1}{2}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}x_{\mu_2}\frac{N^{2\gamma}}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2}\\
&\qquad - \frac{1}{3}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}x_{\mu_2}^3\frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2} - \frac{1}{2}\sum_{\mu_1,\mu_2,\mu_3}^{*} x_{\mu_1}x_{\mu_2}x_{\mu_3}^2\frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2} \qquad (3.4)\\
&\qquad - \frac{1}{12}\sum_{\mu_1,\mu_2,\mu_3,\mu_4}^{*} x_{\mu_1}x_{\mu_2}x_{\mu_3}x_{\mu_4}\frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\Bigg) + \mathcal{O}\big(N|R_N(x/N^\gamma,\xi)|\big),
\end{aligned}
$$

where $\|x\|_4^4 = \sum_{\mu=1}^{M(N)} x_\mu^4$. Here and in the sequel, we use the notation $\sum^*_{\mu_1,\dots,\mu_k}$ for summation over all k-tuples $(\mu_1,\dots,\mu_k) \in \{1,\dots,M(N)\}$ with pairwise disjoint components.

Proof. Just compute the Taylor expansion of $\log(\cosh(\cdot))$. □

The idea now is that all summands that carry a term of the form $\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}$ or $\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}$ will asymptotically vanish ($\mathbb{P}$-almost surely, as $N \to \infty$) due to the law of large numbers. However, since we are integrating with respect to x and x may be growing with N we have to be a bit careful when applying this idea. To control events when one of $\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}$, $\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}$, or $\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\xi_i^{\mu_5}\xi_i^{\mu_6}$ is large for some choice of pairwise different indices μ_i let us prove the following Lemma.

Lemma 3.4. *For $\delta_N = \sqrt{\frac{16\log N}{N}}$ and $\delta'_N = \sqrt{\frac{16\log N}{N^{1-4\gamma}}}$ and $0 < \gamma < 1/4$, define the events*

$$A_{N,\delta'_N}^{\mu_1,\mu_2} = \left\{\left|\frac{1}{N^{1-2\gamma}}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\right| > \delta'_N\right\},$$

$$B_{N,\delta_N}^{\mu_1,\dots,\mu_4} = \left\{\left|\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\right| > \delta_N\right\},$$

$$C_{N,\delta_N}^{\mu_1,\dots,\mu_6} = \left\{\left|\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\xi_i^{\mu_5}\xi_i^{\mu_6}\right| > \delta_N\right\}$$

and

$$D_{N,\delta_N}^{\mu_1,\mu_2} = \left\{\left|\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\right| > \delta_N\right\}.$$

Finally set

$$
\begin{aligned}
\Omega_1(N,\delta_N,\delta'_N) \;&:=\; \Omega_1(N) := \qquad (3.5)\\
&\Bigg(\bigcup_{\mu_1,\mu_2} A_{N,\delta'_N}^{\mu_1,\mu_2} \cup \bigcup_{\mu_1,\dots,\mu_4} B_{N,\delta_N}^{\mu_1,\dots,\mu_4} \cup \bigcup_{\mu_1,\dots,\mu_6} C_{N,\delta_N}^{\mu_1,\dots,\mu_6} \cup \bigcup_{\mu_1,\mu_2} D_{N,\delta_N}^{\mu_1,\mu_2}\Bigg)^c,
\end{aligned}
$$

where each of the unions is taken over all sets of pairwise different indices in $\{1, \ldots, M(N)\}$. *There exists an* $N_1 \in \mathbb{N}$ *and a constant* C *such that for all* $N \geq N_1$

$$\mathbb{P}\big(\Omega_1(N)^c\big) \leq \frac{C}{N^2}. \tag{3.6}$$

Proof. For pairwise different indices $\mu_i \in \{1, \ldots, M(N)\}$, Chebyshev's inequality implies for all $t > 0$

$$\mathbb{P}\big(A^{\mu_1,\mu_2}_{N,\delta'_N}\big) \leq \exp\{-t\, \delta'_N N^{1-2\gamma}\} \exp\{N t^2/2\}.$$

Choosing $t = \delta'_N N^{-2\gamma}$ yields

$$\mathbb{P}\big(A^{\mu_1,\mu_2}_{N,\delta'_N}\big) \leq \exp\left(-N^{1-4\gamma}\frac{(\delta'_N)^2}{2}\right) = \exp\big(-8 \log N\big) = N^{-8}.$$

Similarly, $\mathbb{P}\big(B^{\mu_1,\ldots,\mu_4}_{N,\delta_N}\big) \leq N^{-8}$, as well as $\mathbb{P}\big(C^{\mu_1,\ldots,\mu_6}_{N,\delta_N}\big) \leq N^{-8}$, and $\mathbb{P}\big(A^{\mu_1,\mu_2}_{N,\delta_N}\big) \leq N^{-8}$. Therefore,

$$\mathbb{P}\big(\Omega_1(N)^c\big) \leq \big(2M(N)^2 + M(N)^4 + M(N)^6\big)\, N^{-8} \leq \frac{4}{N^2}$$

for N large enough. □

Lemma 3.4 implies that $(\Omega_1(N))^c$ will only happen finitely often in N with probability one. Let us denote by $(\Omega, \mathcal{F}, \mathbb{P})$ a rich enough probability space where the pattern matrix ξ lives on. Hence there exists a subset $\Omega_1 \subseteq \Omega$ of Ω with $\mathbb{P}(\Omega_1) = 1$ such that for all $\xi \in \Omega_1$ there is a $N(\xi)$ such that for all $N \geq N(\xi)$ we have that

$$\left|\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\right| \leq \delta_N, \qquad \left|\frac{1}{N^{1-2\gamma}}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\right| \leq \delta'_N,$$

as well as

$$\left|\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\right| \leq \delta_N, \quad \text{and} \quad \left|\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\xi_i^{\mu_5}\xi_i^{\mu_6}\right| \leq \delta_N,$$

where δ_N and δ'_N are defined as in Lemma 3.4 (and converge to zero). This will be exploited when studying the behaviour of $-N\Phi_N(x/N^\gamma)$. In fact, the following holds true.

Lemma 3.5. *Define the measure* $\overline{\mathcal{Q}}_{N,\gamma}$ *on* $\mathbb{R}^{M(N)}$ *as the probability measure with density proportional to* $\exp(-N^{1-4\gamma}\,\Psi(x))$, $x \in \mathbb{R}^{M(N)}$. *Here*

$$\Psi(x) := \left(-\frac{1}{12}\|x\|_4^4 - \frac{1}{4}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}^2 x_{\mu_2}^2\right). \tag{3.7}$$

Assume that $M(N)$ *grows so slowly that*

$$\lim_{N\to\infty} \frac{M^6(N)\log N}{N} = 0 \tag{3.8}$$

and

$$\lim_{N\to\infty} \frac{M^2(N)\log N}{N^{1-4\gamma}} = 0. \tag{3.9}$$

Then there exists a set $\tilde{\Omega}(N) \subset \Omega$ with $\mathbb{P}(\tilde{\Omega}(N)) \geq C/N^2$ for some constant $C > 0$ and for all N large enough such that for all $\xi \in \tilde{\Omega}(N)$, for all $k \in \mathbb{N}$ fixed, for all $A_k \in \mathcal{B}(\mathbb{R}^k)$ and

$$A := A_k \times \mathbb{R}^{M(N)-k}$$

it holds that

$$\lim_{N\to\infty} \frac{1}{N^{1-4\gamma}} \left|\log \mathcal{Q}_{N,1,N^{2\gamma}}(A) - \log \overline{\mathcal{Q}}_{N,\gamma}(A)\right| = 0. \tag{3.10}$$

Remark 3.6. We have formulated Lemma 3.5 for sets of the form $A_k \times \mathbb{R}^{M(N)-k}$ since we will prove an MDP for the overlap parameter with respect to $\varrho_{N,\beta}$. Therefore we have to restrict ourselves to finite-dimensional sets to be able to get a well-defined principle. However, the proof of Lemma 3.5 shows that this Lemma might be extended such that (3.10) holds in the supremum taken over a whole class of sets $A \in \mathcal{B}(\mathbb{R}^{M(N)})$.

Remark 3.7. In the above lemma the quality of the approximation of $\mathcal{Q}_{N,1,N^{2\gamma}}(\cdot)$ by $\overline{\mathcal{Q}}_{N,\gamma}(\cdot)$ becomes worse and worse the closer we get to the CLT regime, i.e. the regime of $\gamma = 1/4$. This is reflected in increasingly more restrictive conditions on the number of patterns $M(N)$ (namely we need to have that $M(N)\,\delta'_N = \sqrt{\frac{M(N)^2 \log N}{N^{1-4\gamma}}}$ goes to zero as N goes to infinity). This may come somewhat surprising (even though we will see later that for a regime extremely close to the CLT regime a moderate deviation principle does not hold, not even for a fixed number of patterns). Indeed, using strong Gaussian approximation ideas as in the proof of the main theorem in [16] e.g., one could, in principle improve a bit on this condition by demanding that only

$$\sqrt{\frac{M(N)\log N}{N^{1-4\gamma}}} \to 0$$

as $N \to \infty$, when additionally also

$$\lim_{N\to\infty} \frac{M^{13}(N)}{N} = 0. \tag{3.11}$$

This would indeed slightly improve our conditions if γ is very close to $1/4$. However, we feel that this is only very marginal improvement rather spoils the transparency of the paper and is not really worth the additional expense of several pages of proof.

Before turning to the proof of Lemma 3.5, we quote an estimate which will happen to be very useful in the sequel. It is a bound on the operator norm of the random matrix arising from the patterns. We denote by $\|\cdot\|_{\mathrm{Op}}$ the operator norm on $M(N)\times M(N)$ matrices, defined by $\|A\|_{\mathrm{Op}} = \sup_{\|x\|_2\leq 1} \|Ax\|_2$. Let us recall the notation $\alpha(N) := M(N)/N$.

Lemma 3.8 ([4, Theorem 4.1])**.** *There exist a constant $K > 0$ and a $N_2 \in \mathbb{N}$ such that*

$$\mathbb{P}\left(\left|\left\|\frac{1}{N}\xi^T\xi\right\|_{\mathrm{Op}} - \left(1+\sqrt{\alpha(N)}\right)^2\right| \geq \sqrt{\alpha(N)}\right) \leq K\, e^{-M(N)/K}$$

for all $N \geq N_2$.

For later use, we define the following subset of Ω

$$\Omega_2(N) := \left\{ \xi : \left| \left\| \frac{1}{N} \xi^T \xi \right\|_{\mathrm{Op}} - (1 + \sqrt{\alpha(N)})^2 \right| \leq \sqrt{\alpha(N)} \right\}.$$

In particular, we know that for $N \geq N_2$, $\xi \in \Omega_2(N)$ and all $x, y \in \mathbb{R}^{M(N)}$,

$$\left| \frac{1}{N} \sum_{i=1}^N \langle x, \xi_i \rangle \langle y, \xi_i \rangle - \langle x, y \rangle \right| \leq 4 \sqrt{\alpha(N)} \|x\|_2 \|y\|_2. \tag{3.12}$$

Note that $\mathbb{P}\big(\Omega_2(N)^c\big) \leq C/N^2$, if $M(N)$ is growing faster than some multiple of $\log N$ which we shall assume for convenience.

Proof of Lemma 3.5. Let $A \in \mathcal{B}(\mathbb{R}^{M(N)})$ such that $A = A_k \times \mathbb{R}^{M(N)-k}$ with $A_k \in \mathcal{B}(\mathbb{R}^k)$. In order to investigate $\mathcal{Q}_{N,1,N^{2\gamma}}(A)$ we need to study

$$\frac{\int_A \exp\big(-N \Phi_N(x/N^\gamma)\big)\, dx}{\int_{\mathbb{R}^{M(N)}} \exp\big(-N \Phi_N(x/N^\gamma)\big)\, dx}.$$

We will first concentrate on the numerator and use the Taylor expansion for Φ_N derived in Lemma 3.3. To estimate the different vanishing parts of Φ_N observe that on the set $\Omega_1(N)$ (defined as in Lemma 3.4) and for $N \geq N_1$ we have

$$\begin{aligned}
\left| \frac{1}{2} \sum_{\mu_1, \mu_2}^* x_{\mu_1} x_{\mu_2} \frac{N^{2\gamma}}{N} \sum_{i=1}^N \xi_i^{\mu_1} \xi_i^{\mu_2} \right| &\leq \frac{1}{2} \sqrt{\frac{16 \log N}{N^{1-4\gamma}}} \sum_{\mu_1, \mu_2}^* |x_{\mu_1}|\, |x_{\mu_2}| \\
&\leq \frac{1}{2} \sqrt{\frac{16 \log N}{N^{1-4\gamma}}} \left(\sum_{\mu=1}^M (N) |x_\mu| \right)^2 \\
&\leq \frac{M(N)}{2} \sqrt{\frac{16 \log N}{N^{1-4\gamma}}} \|x\|_2^2.
\end{aligned}$$

For the other ξ-dependent terms we obtain by definition that for all $N \geq N_1$ and all $\xi \in \Omega_1(N)$ (with $\delta_N = \sqrt{\frac{16 \log N}{N}}$)

$$\begin{aligned}
\left| \frac{1}{3} \sum_{\mu_1, \mu_2}^* x_{\mu_1} x_{\mu_2}^3 \frac{1}{N} \sum_{i=1}^N \xi_i^{\mu_1} \xi_i^{\mu_2} \right| &\leq \frac{1}{3} \sum_{\mu_1, \mu_2, \mu_3} |x_{\mu_1} x_{\mu_2}|\, x_{\mu_3}^2 \left| \frac{1}{N} \sum_{i=1}^N \xi_i^{\mu_1} \xi_i^{\mu_2} \right| \\
&= \frac{1}{3} \delta_N \|x\|_2^2 \left(\sum_{\mu=1}^M (N) |x_\mu| \right)^2 \leq \frac{1}{3} M(N) \delta_N \|x\|_2^4
\end{aligned}$$

as well as (following the same ideas)

$$\begin{aligned}
&\left|\frac{1}{2}\sum\nolimits^{*}_{\mu_1,\mu_2,\mu_3} x_{\mu_1}x_{\mu_2}x_{\mu_3}^2\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\right| \\
&\quad = \left|\frac{1}{2}\sum\nolimits^{*}_{\mu_1,\mu_2} x_{\mu_1}x_{\mu_2}\|x\|_2^2\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2} - \sum\nolimits^{*}_{\mu_1,\mu_2} x_{\mu_1}x_{\mu_2}^3\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\right| \\
&\quad \leq \frac{3}{2}M(N)\delta_N\|x\|_2^4.
\end{aligned}$$

Moreover, for $N \geq N_1$ and $\xi \in \Omega_1(N)$, by definition,

$$\begin{aligned}
&\left|\frac{1}{12}\sum\nolimits^{*}_{\mu_1,\mu_2,\mu_3,\mu_4} x_{\mu_1}x_{\mu_2}x_{\mu_3}x_{\mu_4}\frac{1}{N}\sum_{i=1}^N \xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\right| \\
&\quad \leq \frac{\delta_N}{12}\sum_{\mu_1,\mu_2,\mu_3,\mu_4}|x_{\mu_1}x_{\mu_2}x_{\mu_3}x_{\mu_4}| \leq \frac{\delta_N M(N)^2}{12}\|x\|_2^4 = \frac{1}{3}\sqrt{\frac{M(N)^4\log N}{N}}\|x\|_2^4.
\end{aligned}$$

It remains to consider the remainder of the Taylor expansion. Observe that $|h(t)| \leq 2|t|$ and $0 < \theta < 1$. Hence by Cauchy–Schwarz' inequality we obtain

$$|R_N(y,\xi)| \leq \frac{2}{15N}\sum_{i=1}^N\langle y,\xi_i\rangle^6 \leq \frac{2}{15}\sum_{\mu_1,\dots,\mu_6}|y_{\mu_1}\dots y_{\mu_6}|\left|\frac{1}{N}\sum_{i=1}^N\xi_i^{\mu_1}\dots\xi_i^{\mu_6}\right|.$$

The right-hand side is bounded above by a combinatorial factor times the sum of terms similar to the ones treated before (with two, four or six different ξ_i^μ) plus the term arising from $\mu_1 = \dots = \mu_6$. This yields

$$\begin{aligned}
|R_N(y,\xi)| \leq C\Bigg[\sqrt{\frac{M(N)^2\log N}{N}}\|y\|_2^6 &+ \sqrt{\frac{M(N)^4\log N}{N}}\|y\|_2^6 \\
&+ \sqrt{\frac{M(N)^6\log N}{N}}\|y\|_2^6 + \|y\|_6^6\Bigg]
\end{aligned}$$

for $N \geq N_1$ and $\xi \in \Omega_1(N)$. Therefore

$$N|R_N(x/N^\gamma,\xi)| \leq CN^{1-6\gamma}\left[\sqrt{\frac{M(N)^6\log N}{N}}\|x\|_2^6 + \|x\|_6^6\right].$$

Hence, if we assume that $N \geq N_1$ and ξ is in $\Omega_1(N)$ we see that $-N\Phi(x/N^\gamma)$ and $N^{1-4\gamma}\Psi(x)$ differ by at most a constant times $N^{1-4\gamma}$ times

$$\begin{aligned}
r_N(x) = \frac{M(N)}{2}\delta'_N\|x\|_2^2 &+ \left(2\delta_N M(N) + \delta_N M(N)^2\right)\|x\|_2^4 \\
&+ 3N^{-2\gamma}\delta_N M(N)^3\|x\|_2^6 + N^{-2\gamma}\|x\|_6^6
\end{aligned}$$

(where again $\delta'_N = \sqrt{\frac{16\log N}{N^{1-4\gamma}}}$ and $\delta_N = \sqrt{\frac{16\log N}{N}}$).

As usual in this kind of limit theorems we will now split the area of integration into three parts.

The first one is the **inner region**. By definition the inner region is given by

$$\|x\|_2 \le R$$

for some $R > a$, $a > 0$, which we will choose large enough later. Since $\|x\|_6^6 \le \|x\|_2^6$ we obtain for the inner region for N large enough and $\xi \in \Omega_1(N)$

$$r_N(x) \le h_N(R), \tag{3.13}$$

where

$$h_N(R) := \frac{M(N)}{2}\delta'_N R^2 + \left(2\delta_N M(N) + \delta_N M(N)^2\right) R^4 + 3N^{-2\gamma}\delta_N M(N)^3 R^6 + N^{-2\gamma} R^6. \tag{3.14}$$

Now $M(N)\,\delta'_N = 4\sqrt{\frac{M(N)^2 \log N}{N^{1-4\gamma}}} \to 0$ for $N \to \infty$ as well as

$$\delta_N M(N) \le \delta_N M(N)^2 \le \delta_N M(N)^3 = 4\sqrt{\frac{M(N)^6 \log N}{N}} \to 0$$

as $N \to \infty$ by our assumptions. This implies that $r_N(x) \to 0$ for x in the inner region uniformly in x.
The **intermediate region** is characterized by

$$rN^\gamma \ge \|x\|_2 \ge R$$

for some r which we still may choose. Then for x in this intermediate region we have that for N large enough and $\xi \in \Omega_1(N)$

$$r_N(x) \le \frac{M(N)}{2}\delta'_N \|x\|_2^2 + \left(6\delta_N M(N)^3 r^2 + r^2\right)\|x\|_2^4.$$

Hence under the assumptions we made, for all $\varepsilon > 0$ there exists $N_3(r) \in \mathbb{N}$, $N_3(r) \ge N_1$ such that for $N \ge N_3(r)$ we obtain

$$r_N(x) \le \varepsilon \|x\|_2^2 + 2r^2\|x\|_2^4.$$

For $\|x\|_2 \ge R$ trivially $\|x\|_2^4 \ge R^2\|x\|_2^2$, such that by choosing r and $\varepsilon > 0$ small enough, we see that there exists an $N_4(R) \in \mathbb{N}$ such that $N_4(R) \ge N_3(r)$ and

$$\begin{aligned} -N\Phi(x/N^\gamma) &\le N^{1-4\gamma}\left(\Psi(x) + r_N(x)\right) \\ &\le N^{1-4\gamma}\left(-\frac{1}{12}\|x\|_4^4 - \frac{1}{4}\sum_{\mu_1,\mu_2}^* x_{\mu_1}^2 x_{\mu_2}^2 + \varepsilon\|x\|_2^2 + 2r^2\|x\|_2^4\right) \\ &\le N^{1-4\gamma}\left(-\frac{1}{12}\|x\|_4^4 - \frac{1}{12}\left[\|x\|_2^4 - \|x\|_4^4\right] + \varepsilon\|x\|_2^2 + 2r^2\|x\|_2^4\right) \\ &\le -N^{1-4\gamma}\frac{R^2}{24}\|x\|_2^2 \end{aligned} \tag{3.15}$$

holds for all $N \ge N_4(R)$, for all $\xi \in \Omega_1(N)$ and all x from the intermediate region.

Therefore, for N large enough and for $\xi \in \Omega_1(N)$ we obtain for the integral in the intermediate region that

$$\begin{aligned}
\left|\int_{A\cap\{R\le\|x\|_2\le rN^\gamma\}} \exp\bigl(-N\,\Phi(x/N^\gamma)\bigr)\,dx\right| \\
&\le \int_{A\cap\{\|x\|_2\ge R\}} \exp\left(-N^{1-4\gamma}\frac{R^2}{24}\|x\|_2^2\right)\,dx \\
&\le \exp\bigl(-N^{1-4\gamma}R^4/48\bigr)\int_{\mathbb{R}^{M(N)}} \exp\left(-N^{1-4\gamma}\frac{R^2}{48}\|x\|_2^2\right)\,dx \\
&= \exp\bigl(-N^{1-4\gamma}R^4/48\bigr)\left(\frac{48\pi}{R^2N^{1-4\gamma}}\right)^{M(N)/2}.
\end{aligned}$$

This bound will allow us to deduce that the integral over the intermediate region is negligible.

It remains to check that also the integral over the **outer region** vanishes. This integral can be decomposed into two parts. First, we show that there exists an $r_0 > 0$ such that the integral over $B(r_0N^\gamma)^c$ is negligible on the right scale and then, in a second step, we show that this r_0 can be replaced by an arbitrarily small $r > 0$. Indeed, this step is well known from e.g. [15] or [16] where it has been used just on another scale. The principle idea is easy: The regimes $B(rN^\gamma)^c$ for any $r > 0$ belong to the large deviation regime of our overlap parameter and hence have exponentially small probabilities.

For convenience, we denote by $f_{\mathrm{CW}}(\beta)$ the free energy in the Curie-Weiss model at temperature $1/\beta$, i. e.,

$$f_{\mathrm{CW}}(\beta) = -\frac{\beta}{2}z(\beta)^2 + \log\cosh(\beta z(\beta)),$$

where $z(\beta) := z^+(\beta)$. Then,

$$\log\cosh x \le \frac{1}{4\beta}x^2 + \max_{t\in\mathbb{R}}\left(-\frac{1}{4\beta}t^2 + \log\cosh t\right) = \frac{1}{4\beta}x^2 + f_{\mathrm{CW}}(2\beta),$$

which implies in particular for $\beta = 1$ that

$$\begin{aligned}
-N\Phi_N(x/N^\gamma) &= -\frac{1}{2}N^{1-2\gamma}\|x\|_2^2 + \sum_{i=1}^N \log\cosh\langle x/N^\gamma, \xi_i\rangle \\
&\le -\frac{1}{2}N^{1-2\gamma}\|x\|_2^2 + \frac{1}{4}N^{-2\gamma}\sum_{i=1}^N\langle x, \xi_i\rangle^2 + N\ f_{\mathrm{CW}}(2). \qquad (3.16)
\end{aligned}$$

Estimating the sum in (3.16) with the help of the bound (3.12) on the random matrix $\frac{1}{N}\xi^T\xi$, we see that there exist $r_0 > 0$ and $N_5(r_0) \in \mathbb{N}$, $N_5(r_0) \ge N_2$, such that

$$-N\Phi(x/N^\gamma) \le -\frac{1}{6}N^{1-2\gamma}\|x\|_2^2 \qquad (3.17)$$

holds for all x satisfying $\|x\|_2 \ge r_0N^\gamma$, all $N \ge N_5(r_0)$ and all $\xi \in \Omega_2(N)$.

Let us now turn to the regime of $rN^\gamma \le \|x\|_2 \le r_0 N^\gamma$ with an arbitrary $0 < r < r_0$. First note that by an easy application of the triangle inequality and the Cauchy-Schwarz inequality we obtain ($\beta = 1$)

$$\begin{aligned}
\Phi(x/N^\gamma) &= \frac{1}{2}\|x/N^\gamma\|_2^2 - \frac{1}{N}\sum_{i=1}^N \log\cosh(\langle x/N^\gamma, \xi_i\rangle) \\
&\ge \mathbb{E}\Big(\frac{1}{2}\langle x/N^\gamma, \xi_1\rangle^2 - \log\cosh\langle x/N^\gamma, \xi_1\rangle\Big) \\
&\quad - \Big|\frac{1}{N}\sum_{i=1}^N \log\cosh\langle x/N^\gamma, \xi_i\rangle - \mathbb{E}\log\cosh\langle x/N^\gamma, \xi_1\rangle\Big| \\
&\ge \mathbb{E}\Big(\frac{1}{2}\langle x/N^\gamma, \xi_1\rangle^2 - \log\cosh\langle x/N^\gamma, \xi_1\rangle\Big) \qquad (3.18) \\
&\quad - \sup_{\|y\|_2 \le r_0}\Big|\frac{1}{N}\sum_{i=1}^N \log\cosh\langle y, \xi_i\rangle - \mathbb{E}\log\cosh\langle y, \xi_1\rangle\Big|.
\end{aligned}$$

The first term in (3.18) is bounded below by

$$c_{r,r_0} = \inf_{y\,:\, r \le \|y\|_2 \le r_0} \mathbb{E}\,\psi(\langle y, \xi_1\rangle),$$

where

$$\psi(t) = t^2/2 - \log\cosh t, \qquad t \in \mathbb{R}.$$

Observe that $\psi(\cdot)$ attains its unique minimum at $t = 0$. The fact that $\langle y, \xi_1\rangle$ is a (finite) Rademacher average (see [21, Chapter I.4], for instance), implies that

$$\mathbb{P}\big(|\langle y, \xi_1\rangle| \ge \tfrac{1}{8}\|y\|_2\big) > 1/3$$

(cf. [14, Lemma 4.3]). This obviously implies $c_{r,r_0} > 0$ because there is a set of positive $\mathbb{P}$-measure, on which ψ is bounded away from its unique minimum at zero. The second summand in (3.18) becomes small due to the law of large numbers, which in this context often is referred to as self-averaging. It is frequently used in the context of spin glasses. Already the proof of [14, Lemma 4.2] shows that

$$\lim_{N\to\infty} \sup_{\|x\|_2 \le r_0}\Big|\frac{1}{N}\sum_{i=1}^N f(\langle x, \xi_i\rangle) - \mathbb{E}f(\langle x, \xi_1\rangle)\Big| = 0 \qquad (3.19)$$

holds $\mathbb{P}$-almost surely for Lipschitz continuous f. Even more one can obtain effective bounds for large (but fixed) N:

Lemma 3.9 ([14, Lemma 4.2], [16, Lemma 4.8])**.** *There exist a constant $c > 0$ and an $N_6 \in \mathbb{N}$ such that for all $\varepsilon > 0$ and all $N \ge \max\{N_6, 2/\varepsilon^2\}$*

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{\|y\|_2 \le r_0}\Big|\frac{1}{N}\sum_{i=1}^N \log\cosh\langle y, \xi_i\rangle - \mathbb{E}\log\cosh\langle y, \xi_1\rangle\Big| \ge (3 + 2r_0)\varepsilon\Big) \\
&\qquad \le 2\exp\big(M(N)(\log(r_0/\varepsilon) + c)\big)\,\exp\big(-N\varepsilon^2/8\big) + \mathbb{P}(\Omega_2(N)^c).
\end{aligned}$$

For

$$\varepsilon = \frac{c_{r,r_0}}{2(3+2r_0)}$$

and

$$\Omega_3(N,r,r_0) = \left\{\xi : \sup_{\|y\|_2 \le r_0} \left| \frac{1}{N} \sum_{i=1}^{N} \log\cosh\langle y, \xi_i\rangle - \mathbb{E}\log\cosh\langle y, \xi_1\rangle \right| \le \frac{c_{r,r_0}}{2} \right\} \quad (3.20)$$

this readily implies that there exist a constant $K(r, r_0) > 0$ and an $N_6(r, r_0) \in \mathbb{N}$ such that for all $N \ge N_6(r, r_0)$

$$\mathbb{P}(\Omega_3(N,r,r_0)^c) \le \exp\bigl(-K(r,r_0)N\bigr) + \mathbb{P}(\Omega_2(N)^c).$$

Our estimates on the two summands on the right-hand side of (3.18) yield

$$-N\Phi(x/N^\gamma) \le -Nc_{r,r_0}/2 \quad (3.21)$$

for all x such that $rN^\gamma \le \|x\|_2 \le r_0 N^\gamma$, all $N \ge N_6(r, r_0)$ and all $\xi \in \Omega_3(N, r, r_0)$.

Now we are ready to assemble all our estimates in order to estimate our desired integral. First of all let us assume that

$$\xi \in \tilde{\Omega}(N) : \;=\; \Omega_1(\delta_N, \delta'_N, N) \cap \Omega_2(N) \cap \Omega_3(N, r, r_0) \quad (3.22)$$

and that

$$N \ge \max\{N_1, N_2, N_3(r), N_4(R), N_5(r_0), N_6(r, r_0)\}.$$

Note that there exists a constant $C > 0$ such that

$$\mathbb{P}(\tilde{\Omega}(N)^c) \le C/N^2. \quad (3.23)$$

Naturally, L depends on our choice of R, r and r_0.
According to what we have seen there is some function $h_N(R)$ (see (3.14)) with $h_N(R) \to 0$ as $N \to \infty$ such that

$$\begin{aligned}
&\int_A \exp\bigl(-N\Phi(x/N^\gamma)\bigr)\,dx \\
&\quad= \exp\bigl(\mathcal{O}(N^{1-4\gamma}h_N(R))\bigr) \int_{A\cap B(0,R)} \exp\bigl(N^{1-4\gamma}\Psi(x)\bigr)\,dx \\
&\qquad+\mathcal{O}\left(\exp\bigl(-N^{1-4\gamma}R^4/48\bigr)\left(\frac{48\pi}{R^2N^{1-4\gamma}}\right)^{M(N)/2}\right) \\
&\qquad+\mathcal{O}\left(\bigl[\exp(-Nr_0^2/12) + \exp(-Nc_{r,r_0}/4)\bigr]\right). \qquad (3.24)
\end{aligned}$$

The last summand in (3.24) vanishes on our (logarithmic) $N^{1-4\gamma}$-scale. The second summand can be made arbitrarily small (on the same scale) by choosing R large. Hence our main task is to replace the first integral

$$\int_{A\cap B(0,R)} \exp\bigl(N^{1-4\gamma}\Psi(x)\bigr)\,dx$$

by the integral over the entire set A. By the same techniques as above one shows that for all x with $\|x\|_2 \geq R$:

$$\Psi(x) \leq -\frac{1}{12}\|x\|_4^4 - \frac{1}{12}\left[\|x\|_2^4 - \|x\|_4^4\right] \leq -\frac{R^2}{24}\|x\|_2^2.$$

Consequently,

$$\begin{aligned}\int_{\{\|x\|_2 \geq R\}} & \exp\bigl(N^{1-4\gamma}\Psi(x)\bigr)dx \\ & \leq \int_{\{\|x\|_2 \geq R\}} \exp\left(-N^{1-4\gamma}\frac{R^2}{24}\|x\|_2^2\right) dx \\ & \leq \exp\bigl(-N^{1-4\gamma}R^4/48\bigr)\left(\frac{48\pi}{R^2 N^{1-4\gamma}}\right)^{M(N)/2}.\end{aligned}$$

By (3.24) this yields that

$$\begin{aligned}\int_A & \exp\bigl(-N\Phi(x/N^\gamma)\bigr)\, dx \\ & = \exp\bigl(\mathcal{O}(N^{1-4\gamma}\, h_N(R)\bigr) \int_A \exp\bigl(N^{1-4\gamma}\Psi(x)\bigr)\, dx \\ & \quad + \mathcal{O}\left(\exp\bigl(-N^{1-4\gamma}R^4/48\bigr)\left(\frac{48\pi}{R^2 N^{1-4\gamma}}\right)^{M(N)/2}\right) \\ & \quad + \mathcal{O}\left(\left[\exp\bigl(-Nr_0^2/12\bigr) + \exp\bigl(-Nc_{r,r_0}/4\bigr)\right]\right).\end{aligned}$$

In order to compare $\mathcal{Q}_{N,1,N^{2\gamma}}(A)$ to $\overline{\mathcal{Q}}_{N,\gamma}(A)$, i.e.

$$\int_A \exp\bigl(-N\Phi(x/N^\gamma)\bigr)\, dx \Big/ \int_{\mathbb{R}^{M(N)}} \exp\bigl(-N\Phi(x/N^\gamma)\bigr)\, dx$$

and

$$\int_A \exp\bigl(N^{1-4\gamma}\Psi(x)\bigr)\, dx \Big/ \int_{\mathbb{R}^{M(N)}} \exp\bigl(N^{1-4\gamma}\Psi(x)\bigr)\, dx,$$

we also need to show that the denominators of the above integrals are asymptotically equivalent on the logarithmic $N^{1-4\gamma}$-scale.

This can be done analogously to the above. However, one will in this case also need a lower bound for $\int_{\mathbb{R}^{M(N)}} \exp(\Psi(x))\, dx$. This is obtained by bounding Ψ from below. As above we estimate

$$\Psi(x) = -\frac{1}{12}\|x\|_4^4 - \frac{1}{4}\left[\|x\|_2^4 - \|x\|_4^4\right] \geq -\frac{1}{4}\|x\|_2^4.$$

For $\|x\|_2 \leq R$ we obtain the bound

$$\Psi(x) \geq -\frac{R^2}{3}\|x\|_2^2.$$

Now,

$$\begin{aligned}\int_A \exp\big(N^{1-4\gamma}\Psi(x)\big)\,dx &\geq \int_{A\cap B(0,R)} \exp\Big(-N^{1-4\gamma}\frac{R^2}{3}\|x\|_2^2\Big)\,dx \\ &\geq \frac{1}{2}\Big(\frac{3\pi}{R^2N^{1-4\gamma}}\Big)^{M(N)/2}.\end{aligned}$$

With these preparations one easily concludes the proof. Choose R_N increasing in N such that $h_N(R_N)$ still goes to zero as N goes to ∞. This is possible since

$$h_N(R_N) = \frac{M(N)}{2}\delta_N' R_N^2 + \big(\delta_N + \delta_N M(N)^2\big)\,R_N^4 + 3N^{-2\gamma}\delta_N M(N)^3 R_N^6.$$

Then

$$\begin{aligned}&\limsup_{N\to\infty}\frac{1}{N^{1-4\gamma}}\left|\log\int_A \exp\big(N^{1-4\gamma}\Psi(x)\big)\,dx - \log\int_A \exp\big(-N\,\Phi(x/N^\gamma)\big)\,dx\right| \\ &\quad= \limsup_{N\to\infty}\frac{1}{N^{1-4\gamma}}\log e^{N^{1-4\gamma}h_N(R_N)} = 0.\end{aligned}$$

A similar argument applies to the denominators of the two integrals defining the measures $\mathcal{Q}_{N,\gamma}(\cdot)$ and $\overline{\mathcal{Q}}_{N,1,N^{2\gamma}}(\cdot)$. This finishes the proof of Lemma 3.5. □

The previous Lemma 3.5 tells us that – in order to study the moderate deviations of the overlap parameter m_N – we only need to study the asymptotic behavior of the normalized integrals

$$\int_A e^{N^{1-4\gamma}\Psi(x)}dx \Big/ \int_{\mathbb{R}^{M(N)}} e^{N^{1-4\gamma}\Psi(x)}dx.$$

This is done in the following lemma.

Lemma 3.10. *Let $k \leq M(N)$ be a fixed number. Let $B^k(r)$ denote the ball of radius $r > 0$ in the first k coordinates centered in the origin, i.e.*

$$B^k(r) := \{x \in \mathbb{R}^{M(N)} : x_1^2 + \ldots + x_k^2 \leq r^2\}.$$

Then under the assumptions of Lemma 3.5 for all $k \in \mathbb{N}$ (fixed) and all $r > 0$ it holds that

$$\lim_{N\to\infty}\frac{1}{N^{1-4\gamma}}\log\int_{(B^k(r))^c}\left(e^{N^{1-4\gamma}\Psi(x)}dx \Big/ \int_{\mathbb{R}^{M(N)}} e^{N^{1-4\gamma}\Psi(x)}\,dx\right) = \sup_{x\notin B^k(r)}\Psi(x). \tag{3.25}$$

The same formula is true when $(B^k(r))^c$ is replaced by $(A\times\mathbb{R}^{M(N)-k})\cap(B^k(r))^c$ for $A\in\mathcal{B}(\mathbb{R}^k)$.

Proof. First we prove the upper bound. According to what we have shown in the proof of Lemma 3.5 we can find a sequence $(R_N)_{N\in\mathbb{N}}$ growing to infinity (slowly enough) such that

$$\limsup_{N\to\infty}\Bigg|\frac{1}{N^{1-4\gamma}}\log\int_{(B^k(r))^c}\left(e^{N^{1-4\gamma}\,\Psi(x)}\,dx\Big/\int_{\mathbb{R}^{M(N)}}e^{N^{1-4\gamma}\,\Psi(x)}dx\right)$$
$$-\frac{1}{N^{1-4\gamma}}\log\int_{B(R_N)\cap(B^k(r))^c}\left(e^{N^{1-4\gamma}\Psi(x)}dx\Big/\int_{\mathbb{R}^{M(N)}}e^{N^{1-4\gamma}\Psi(x)}dx\right)\Bigg|=0.$$

The smaller of these integrals can be estimated by

$$\int_{B(R_N)\cap(B^k(r))^c}e^{N^{1-4\gamma}\Psi(x)}dx\le\lambda^{M(N)}\big(B(R_N)\big)\sup_{x\in(B^k(r))^c}e^{N^{1-4\gamma}\Psi(x)},$$

where $\lambda^{M(N)}(\cdot)$ denotes the $M(N)$-dimensional Lebesgue measure.
Now for an appropriate constant $C>0$

$$\lambda^{M(N)}\big(B(R_N)\big)\big)\le C\,R_N^{M(N)}.$$

This yields

$$\frac{1}{N^{1-4\gamma}}\log\int_{B(R_N)\cap(B^k(r))^c}e^{N^{1-4\gamma}\Psi(x)}dx$$
$$\le\frac{1}{N^{1-4\gamma}}\log C+\frac{M(N)}{N^{1-4\gamma}}\log R_N+\sup_{x\in(B^k(r))^c}\Psi(x).$$

Now clearly $\frac{1}{N^{1-4\gamma}}\log C\to 0$ for $N\to\infty$ and also

$$\frac{M(N)}{N^{1-4\gamma}}\log R_N\le\frac{M(N)^2\log N}{N^{1-4\gamma}}\to 0$$

for $N\to\infty$, if we choose R_N appropriately. Therefore

$$\limsup_{N\to\infty}\frac{1}{N^{1-4\gamma}}\log\int_{B(R_N)\cap(B^k(r))^c}e^{N^{1-4\gamma}\Psi(x)}dx\le\sup_{x\in(B^k(r))^c}\Psi(x),$$

which is the upper bound.

For the lower bound, observe that $\Psi(\cdot)$ is a continuous function. Even more, for $\varepsilon>0$, if $\|x-y\|<\varepsilon^{1/4}/N^{1/4-\gamma}$, then also $\|\Psi(x)-\Psi(y)\|<\varepsilon/N^{1-4\gamma}$. For $\varepsilon>0$ given, take any $x\notin B^k(r)$ with

$$\sup_{y\notin B^k(r)}\Psi(y)-\Psi(x)\le\varepsilon/N^{1-4\gamma}.$$

Then for N large enough at least one half of the ball $B(x, \varepsilon^{1/4}/N^{1/4-\gamma})$ is contained in $(B^k(r))^c$. Therefore

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \int_{(B^k(r))^c} e^{N^{1-4\gamma}\Psi(x)} dx \\
&\geq \liminf_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \int_{(B^k(r))^c \cap B(x,\varepsilon^{1/4}/N^{1/4-\gamma})} e^{N^{1-4\gamma}\Psi(x)} dx \\
&\geq \liminf_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \Big[\exp\Big(N^{1-4\gamma} (\sup_{y\notin B^k(r)} \Psi(y) - 2\varepsilon/N^{1-4\gamma}) \Big) \times \\
&\lambda^{M(N)}(B(x, \varepsilon^{1/4}/N^{1/4-\gamma})) \Big].
\end{aligned}$$

Now

$$\lambda^{M(N)}(B(x, \varepsilon^{1/4}/N^{1/4-g})) = v_{M(N)} \frac{\varepsilon^{M(N)/4}}{N^{M(N)/4 - M(N)\gamma}},$$

where

$$v_{M(N)} = \lambda^{M(N)}(B(1))$$

is the volume of the $M(N)$-dimensional unit ball. The asymptotics for

$$v_{M(N)} = \begin{cases} \frac{1}{(M(N)/2)!} \pi^{M(N)/2}, & \text{if } M(N) \text{ is even}, \\ \frac{2^{(M(N)+1)/2}}{(M(N)/2-1)!!} \pi^{(M(N)-1)/2}, & \text{if } M(N) \text{ is odd}, \end{cases}$$

together with Stirling's formula for the factorials yields

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \lambda^{M(N)}(B(x, \varepsilon^{1/4}/N^{1/4-\gamma})) \\
&\geq \liminf_{N\to\infty} \left[\frac{M(N)}{N^{1-4\gamma}} \left(\frac{1}{4} \log \varepsilon - (1/4 - \gamma) \log N + \log \pi + \log M(N) - 1 \right) \right] = 0.
\end{aligned}$$

This implies

$$\liminf_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \int_{(B^k(r))^c} e^{N^{1-4\gamma}\Psi(x)} dx \geq \sup_{y\notin B^k(r)} \Psi(y) - 2\varepsilon.$$

Since this is true for all $\varepsilon > 0$, we obtain our desired result. □

Now we are ready to prove the central result of this section, the MDP at the critical temperature.

Proof of Theorem 1.2. According to Lemma 3.2 for each fixed $k \in \mathbb{N}$ the sequences of variables $\Pi_k(N^\gamma m_N)$ and $\Pi_k(N^\gamma m_N) + Y_N$ have identical moderate deviations behavior $\mathbb{P}$-almost surely. According to Lemma 3.1 the latter variable has distribution $\mathcal{Q}_{N,1,N^{2\gamma}}$. In particular this distribution is absolutely continuous with respect to $\lambda^{M(N)}$ and the density is given by the (random) function $f_{N,1,N^{2\gamma}}$ defined in (3.1). Let $k \in \mathbb{N}$ be fixed and let $A \in \mathcal{B}(\mathbb{R}^k)$. Now,

$$\varrho_{N,1}\big(\Pi_k(N^\gamma m_N) \in A\big) = \varrho_{N,1}\big(N^\gamma m_N \in A \times \mathbb{R}^{M(N)-k}\big)$$

and according to Lemma 3.5 we obtain for all $A \in \mathcal{B}(\mathbb{R}^k)$

$$\lim_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \mathcal{Q}_{N,1,N^{2\gamma}}\big(A \times \mathbb{R}^{M(N)-k}\big) = \lim_{N\to\infty} \frac{1}{N^{1-4\gamma}} \log \overline{\mathcal{Q}}_{N,\gamma}\big(A \times \mathbb{R}^{M(N)-k}\big).$$

The aim is to prove a MDP for a sequence $(X_N)_{N\in\mathbb{N}}$ distributed according to $\mathcal{Q}_{N,1,N^{2\gamma}}$. Therefore consider a arbitrary closed subset $A \in \mathcal{B}(\mathbb{R}^k)$. For each $N \in \mathbb{N}$ there exists an $\varepsilon_N > 0$ with $\lim_{N\to\infty} \varepsilon_N = 0$ such that

$$\mathcal{Q}_{N,1,N^{2\gamma}}\big(A \times \mathbb{R}^{M(N)-k}\big) \le \overline{\mathcal{Q}}_{N,\gamma}\big(A \times \mathbb{R}^{M(N)-k}\big) + e^{N^{1-4\gamma}\varepsilon_N}.$$

If $0 \in A$ then on the one hand $\lim_{N\to\infty} \overline{\mathcal{Q}}_{N,\gamma}(A \times \mathbb{R}^{M(N)-k}) > 0$ and on the other hand $\inf_{x\in A\times\mathbb{R}^{M(N)-k}} \Psi(x) = \Psi(0) = 0$. This show that we can concentrate on a set $(A \times \mathbb{R}^{M(N)-k}) \cap (B^k(r))^c$ for some $r > 0$. Since (3.25) is also true when the set $(B^k(r))^c$ is replaced by $(A \times \mathbb{R}^{M(N)-k}) \cap (B^k(r))^c$, we end with

$$\limsup_{N\to\infty} \frac{1}{N^{1-4\gamma}} Q_{N,1,N^{2\gamma}}\big(A \times \mathbb{R}^{M(N)-k}\big) \le \sup_{x\in A\times\mathbb{R}^{M(N)-k}} \Psi(x).$$

The lower bound can be obtained by the same arguments. For an arbitrary open subset $A \in \mathcal{B}(\mathbb{R}^k)$, take a ball $B(x,R) \subset A$, so that $Q_{N,1,N^{2\gamma}}\big(A \times \mathbb{R}^{M(N)-k}\big) \ge Q_{N,1,N^{2\gamma}}\big(B(x,R) \times \mathbb{R}^{M(N)-k}\big)$ and apply (3.25). □

Proof of Theorem 1.4. As in the proof of Lemma 3.2 we see that at $\beta = 1$ and on the right scale (which in our case is easily computed as b_N^4) the moderate deviations behavior of $\varrho_{N,\beta} \circ \big(\Pi_k(\frac{N^{1/4}}{b_N} m_N)\big)^{-1}$ is governed by that of $\mathcal{Q}_{N,1,a_N}$ where we choose $a_N = N^{1/2}/b_N^2$. Now, as in Lemma 3.3 we can taylorexpand the exponent in the density of $\mathcal{Q}_{N,1,a_N}$. Now even, if $M(N)$ is growing so slowly that all others terms vanish nicely ($\mathbb{P}$-almost surely), in the limit as $N \to \infty$, we will in any case meet the terms

$$b_N^4 \left(-\frac{1}{12}\|x\|_4^4 - \frac{1}{4}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}^2 x_{\mu_2}^2 + \frac{1}{2}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1} x_{\mu_2} \frac{1}{\sqrt{N b_N^4}} \sum_{i=1}^{N} \xi_i^{\mu_1}\xi_i^{\mu_2} \right).$$

Now the first two summands do not really matter, but the third summand for $\mathbb{P}$-almost all realizations of ξ does not admit a limit, i.e.

$$\limsup_{N\to\infty} \frac{1}{\sqrt{N b_N^4}} \sum_{i=1}^{N} \xi_i^{\mu_1}\xi_i^{\mu_2} = \infty$$

and

$$\liminf_{N\to\infty} \frac{1}{\sqrt{N b_N^4}} \sum_{i=1}^{N} \xi_i^{\mu_1}\xi_i^{\mu_2} = -\infty.$$

Therefore the last term in the expansion of the exponent above will only have a limit for the third summand, if the x_μ are zero. Hence an MDP cannot exist. □

Proof of Theorem 1.6. We only give a short sketch of the proof. Let $\delta_N = N^{2\gamma}(\beta_N - 1)$. Then $\delta_N \to 0$ corresponds to the case $|\beta_N - 1| \ll 1/N^{2\gamma}$, where we expect the same asymptotic behaviour as for $\beta_N = 1$, whereas $\delta_N = \delta \neq 0$ corresponds to the borderline case $\beta_N - 1 = \delta\,(1/N^{2\gamma})$. With the Taylor expansion of the function $\Phi_N(x)$ given in lemma 3.3 we find for every $\beta_N > 0$:

$$
\begin{aligned}
&-N\,\Phi_N(x/N^\gamma)\\
&\quad= -N^{1-2\gamma}(1-\beta_N)\frac{\beta_N}{2}\|x\|_2^2 - N^{1-4\gamma}\Bigg(-\frac{\beta_N^4}{12}\|x\|_4^4 - \frac{\beta_N^4}{4}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}^2 x_{\mu_2}^2\\
&\qquad+\frac{\beta_N^2}{2}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}x_{\mu_2}\frac{N^{2\gamma}}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2} - \frac{\beta_N^4}{3}\sum_{\mu_1,\mu_2}^{*} x_{\mu_1}x_{\mu_2}^3\frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2}\\
&\qquad-\frac{\beta_N^4}{2}\sum_{\mu_1,\mu_2,\mu_3}^{*} x_{\mu_1}x_{\mu_2}x_{\mu_3}^2\frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2} \qquad\qquad (3.26)\\
&\qquad-\frac{\beta_N^4}{12}\sum_{\mu_1,\mu_2,\mu_3,\mu_4}^{*} x_{\mu_1}x_{\mu_2}x_{\mu_3}x_{\mu_4}\frac{1}{N}\sum_{i=1}^{N}\xi_i^{\mu_1}\xi_i^{\mu_2}\xi_i^{\mu_3}\xi_i^{\mu_4}\Bigg) + \mathcal{O}\big(N|R_{N,\beta_N}(x/N^\gamma,\xi)|\big).
\end{aligned}
$$

If we consider the case that $|\beta_N - 1| \ll 1/N^{2\gamma}$, we observe that $N^{4\gamma-1}\,N^{1-2\gamma}(1-\beta_N) = N^{2\gamma}(1-\beta_N) \to 0$ as $N \to \infty$. Hence considering the scale (speed) $N^{1-4\gamma}$, the first term in the Taylor expansion (3.26), $-N^{1-2\gamma}(1-\beta_N)\frac{\beta_N}{2}\|x\|_2^2$, does not appear. So we obtain the MDP along the lines of the proof of Theorem 1.2 and get the same rate function.

In the second case, we assume that $|\beta_N - 1|\,N^{2\gamma} \to \infty$ for $N \to \infty$. Now with speed $|1-\beta_N|\,N^{1-2\gamma}$ the first term in (3.26) converges to $(1/2)\|x\|_2^2$, whereas every higher order term in the expansion vanishes, using $N^{1-4\gamma}/(N^{1-2\gamma}|1-\beta_N|) = 1/(N^{2\gamma}|1-\beta_N|) \to 0$ if $N \to \infty$. Along the lines of the proof of Theorem 1.2 we see, that we end with the non-critical rate function $(1/2)\sum_{i=1}^k x_i^2$.

Finally we consider the case $|\beta_N - 1| = \delta(1/N^{2\gamma})$ for some $\delta > 0$. Now on the speed-scale $N^{1-4\gamma}$ the first in the Taylor expansion (3.26), $-N^{1-2\gamma}(1-\beta_N)\frac{\beta_N}{2}\|x\|_2^2$, converges to $\delta\frac{1}{2}\|x\|_2^2$. The higher order terms behave like in the proof of Theorem 1.2, so the rate function is the super-position of the Gaussian and the critical rate function. ☐

References

1. A. Bovier and V. Gayrard, *An almost sure large deviation principle for the Hopfield model*, Ann. Probab. **24** (1996), no. 3, 1444–1475. MR **98g:**60055
2. ———, *An almost sure central limit theorem for the Hopfield model*, Markov Process. Related Fields **3** (1997), no. 2, 151–173. MR **98j:**60044
3. ———, *The retrieval phase of the Hopfield model: a rigorous analysis of the overlap distribution*, Probab. Theory Related Fields **107** (1997), no. 1, 61–98. MR **98b:**82041
4. A. Bovier and V. Gayrard, *Hopfield models as generalized random mean field models*, Mathematical aspects of spin glasses and neural networks, Progr. Probab., vol. 41, Birkhäuser Boston, Boston, MA, 1998, pp. 3–89. MR **99a:**82049

5. A. Bovier, V. Gayrard, and P. Picco, *Gibbs states of the Hopfield model in the regime of perfect memory*, Probab. Theory Related Fields **100** (1994), no. 3, 329–363. MR **96c:**60115
6. A. Bovier and D. M. Mason, *Extreme value behavior in the Hopfield model*, Ann. Appl. Probab. **11** (2001), no. 1, 91–120. MR **2002c:**82049
7. A. Bovier and P. Picco (eds.), *Mathematical aspects of spin glasses and neural networks*, Progress in Probability, vol. 41, Birkhäuser Boston Inc., Boston, MA, 1998. MR **98i:**82002
8. H. J. Brascamp and E. H. Lieb, *On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation*, J. Functional Analysis **22** (1976), no. 4, 366–389. MR 56 #8774
9. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, New York, 1998.
10. J.-D. Deuschel, G. Giacomin, and D. Ioffe, *Large deviations and concentration properties for $\nabla\phi$ interface models*, Probab. Theory Related Fields **117** (2000), no. 1, 49–111. MR **2001e:**60206
11. P. Eichelsbacher and M. Löwe, *Moderate deviations for a class of mean-field models*, to appear in MPRF, 2004.
12. R. S. Ellis and C. M. Newman, *Limit theorems for sums of dependent random variables occurring in statistical mechanics*, Z. Wahrsch. Verw. Gebiete **44** (1978), no. 2, 117–139. MR 58 20113
13. B. Gentz, *An almost sure central limit theorem for the overlap parameters in the Hopfield model*, Stochastic Process. Appl. **62** (1996), no. 2, 243–262. MR **97e:**60038
14. ———, *A central limit theorem for the overlap in the Hopfield model*, Ann. Probab. **24** (1996), no. 4, 1809–1841. MR **97j:**60038
15. ———, *On the central limit theorem for the overlap in the Hopfield model*, Mathematical aspects of spin glasses and neural networks, Birkhäuser Boston, Boston, MA, 1998, pp. 115–149. MR **98m:**82059
16. B. Gentz and M. Löwe, *Fluctuations in the Hopfield model at the critical temperature*, Markov Process. Related Fields **5** (1999), no. 4, 423–449. MR **2001k:**60140
17. ———, *The fluctuations of the overlap in the Hopfield model with finitely many patterns at the critical temperature*, Probab. Theory Related Fields **115** (1999), no. 3, 357–381. MR **2000m:**60111
18. J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proc. Natl. Acad. Sci. U.S.A. **79** (1982), 2554–2558.
19. J. Komlós and R. Paturi, *Convergence results in an associative memory model*, Neural Networks **1** (1988), 239–250.
20. C. Külske, *Metastates in disordered mean-field models: random field and Hopfield models*, J. Statist. Phys. **88** (1997), no. 5-6, 1257–1293. MR **2000a:**82038
21. M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer-Verlag, Berlin, 1991.
22. D. Loukianova, *Lower bounds on the restitution error in the Hopfield model*, Probab. Theory Related Fields **107** (1997), 161–176.
23. M. Löwe, *On the storage capacity of Hopfield models with correlated patterns*, Ann. Appl. Probab. **8** (1998), no. 4, 1216–1250. MR **99k:**82059
24. R. J. McEliece, E. C. Posner, E. R. Rodemich, and S. S. Venkatesh, *The capacity of the Hopfield associative memory*, IEEE Trans. Inform. Theory **33** (1987), 461–482.
25. C.M. Newman, *Memory capacity in neural network models: Rigorous lower bounds*, Neural Networks **1** (1988), 223–238.
26. L. Pastur, M. Shcherbina, and B. Tirozzi, *On the replica symmetric equations for the Hopfield model*, J. Math. Phys. **40** (1999), no. 8, 3930–3947. MR **2001a:**82045
27. L. A. Pastur and A. L. Figotin, *Exactly soluble model of a spin glass*, Sov. J. Low Temp. Phys. **3** (1977), no. 6, 378–383.
28. ———, *On the theory of disordered spin systems*, Theor. Math. Phys. **35** (1977), 403–414.
29. D. Sherrington and S. Kirkpatrick, *Solvable model of a spin–glass*, Phys. Rev. Lett. **35** (1975), 1796–1972. MR 31 4107

30. M. Talagrand, *Rigorous results for the Hopfield model with many patterns*, Probab. Theory Related Fields **110** (1998), no. 2, 177–276. MR **98k:**82099
31. ______, *Rigorous low-temperature results for the mean field p-spins interaction model*, Probab. Theory Related Fields **117** (2000), no. 3, 303–360. MR **2001f:**82037
32. ______, *On the Hopfield model at the critical temperature*, Probab. Theory Related Fields **121** (2001), no. 2, 237–268. MR 1 865 487

(Peter Eichelsbacher) Ruhr-Universität Bochum, Fakultät für Mathematik, NA3/68, 44780 Bochum, Germany
E-mail address, Peter Eichelsbacher: peter.eichelsbacher@ruhr-uni-bochum.de

(Matthias Löwe) Institut für Mathematische Statistik, Fachbereich Mathematik und Informatik, Universität Münster, Einsteinstr. 62, 48149 Münster, Germany
E-mail address, Matthias Löwe: maloewe@math.uni-muenster.de