Increasing propagation of chaos for mean field models

G. Ben Arous *
DMI, Ecole Normale Superieure
45 Rue d'Ulm
75230 Paris Cedex 05, France

O. Zeitouni †
Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

November 17, 1996. Revised January 27, 1998 and June 16, 1998

AMS Subject classification: Primary 60F10. Secondary 60G09, 60K35.

Keywords: Large deviations. Gibbs potential. Exchangeability.

Abstract Let $\mu^{(N)}$ denote a mean-field measure with potential F. Asymptotic independence properties of the measure $\mu^{(N)}$ are investigated. In particular, with $H(\cdot|\mu)$ denoting relative entropy, if there exists a unique non–degenerate minimum of $H(\cdot|\mu) - F(\cdot)$, then propagation of chaos holds for blocks of size $o(N)$. Certain degenerate situations are also studied. The results are applied for the Langevin dynamics of a system of interacting particles leading to a McKean-Vlasov limit.

Résumé Soit $\mu^{(N)}$ une mesure de type champ-moyen avec potentiel d'interaction F. Les propriétés asymptotiques d'indépendance de la mesure $\mu^{(N)}$ sont étudieés. En particulier, si $H(\cdot|\mu)$ designe l'entropie relative, on montre que, s'il existe un unique minimum non dégénéré de $H(\cdot|\mu) - F(\cdot)$, alors la propagation du chaos est valide pour les block de taille $o(N)$. Certains cas de minima dégénéré sont aussi étudiés. Les resultats sont appliqués à la dynamique de Langevin d'un système de particules convergeant vers une limite de McKean-Vlasov.

1 Introduction

Let $(S, \mathcal{S}, \mu)$ denote a (Polish) measure space, let $F : M_1(S) \to \mathbb{R}$ be measurable and bounded, and let $\mathbf{X} = X_1, \ldots, X_N$ denote a sequence of random variables distributed according to the *mean field Gibbs*

*Current address: Dept. of Mathematics, Ecole Polytechnique Fédérale, CH-1015 Lausanne, Suisse

†The work of this author was partially supported by a US-Israel BSF grant.

measure with potential F

$$\mu^{(N)}(dx_1,\ldots,dx_N) = Z_N^{-1}\exp(NF(L_N^{\mathbf{x}}))\prod_{i=1}^N \mu(dx_i)\,. \tag{1}$$

Here $L_N^{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ is the empirical measure of the vector $\mathbf{x} = \{x_i\}_{i=1}^N$, and $Z_N = \int_{S^N}\exp(NF(L_N^{\mathbf{x}}))$ $\prod_{i=1}^N \mu(dx_i)$ is a normalization constant. In various places, we also use $L_N^{\mathbf{X}} := N^{-1}\sum_{i=1}^N \delta_{X_i}$ to denote the (random) empirical measure of the random sample $\mathbf{X}$.

Under mild assumptions, the law of $L_N^{\mathbf{X}}$ converges in distribution, under the law $\mu^{(N)}$, to a deterministic measure δ_{μ^*}. By exchangeability, this implies the convergence of the law of X_1 under $\mu^{(N)}$ to μ^*, and more generaly, for any k finite, the convergence of the law of $(X_1,\ldots,X_k)$ under $\mu^{(N)}$ to $(\mu^*)^{\otimes k}$.

For any exchangeable measure $\theta^{(N)}$ on S^N, let $\theta^{N,k}$ denote its marginal on the first k coordinates; that is, for $A \subset S^k$ measurable,

$$\int_A \theta^{N,k}(dx_1,\ldots,dx_k) = \int_{A\times S^{N-k}} \theta^{(N)}(dx_1,\ldots,dx_N)\,.$$

By exchangeability, for any permutation $\sigma : \{1,\ldots,N\} \to \{1,\ldots,N\}$,

$$\int_{\{(x_1,\ldots,x_N):(x_{\sigma_1},\ldots,x_{\sigma_k})\in A\}} \mu^{(N)}(dx_1,\ldots,dx_N) = \int_A \mu^{N,k}(dx_1,\ldots,dx_k)\,,$$

i.e. the k-marginals of $\mu^{(N)}$ do not depend on the choice of coordinates.

Recall that the relative entropy $H(\cdot|\cdot)$ is defined as

$$H(\nu|\mu) = \begin{cases} \int d\nu \log\frac{d\nu}{d\mu} & \text{if } \nu \ll \mu \\ \infty & \text{otherwise.}\end{cases}$$

In this article, we estimate the relative entropy distance between $\mu^{(N)}$ and appropriate (simpler) exchangeable measures $\nu^{(N)}$, which are related to the law μ^*. The main interest in obtaining such estimates stems from the fact that if one has that

$$B_N = H(\mu^{(N)}|\nu^{(N)})\,, \tag{2}$$

then, for $\nu^{(N)} = \nu^{\otimes N}$ a product measure, and $k(N) = o(N/B_N)$,

$$H(\mu^{N,k(N)}|\nu^{\otimes k(N)}) \le \frac{k(N)}{N} H(\mu^{(N)}|\nu^{\otimes N}) \le \frac{k(N)B_N}{N} \underset{N\to\infty}{\longrightarrow} 0\,. \tag{3}$$

(See e.g. [5, (2.10)] for the first inequality). Hence, for appropriate (large, increasing) blocks $k(N)$, the relative entropy distance (and hence, also the variational distance) between $\mu^{N,k(N)}$ and $\nu^{\otimes k(N)}$ converges to 0. This implies a strong version of the propagation of chaos. It is important to notice that the notion of convergence we use is well suited to deal with increasing blocks: a statement in the weak topology of $M_1(S^{\mathbb{Z}})$ would not be an advance over the finite dimensional propagation of chaos.

We remark that one can, by consideration of the function $F(\nu) = g(\int x d\nu)$ with smooth $g(\cdot)$, provided the latter integral is well defined and that the support of μ is bounded, adapt the set-up described above to Gibbs measures involving empirical means, as opposed to empirical measures. See the remark at the end of Section 2.

Similarly, the ideas presented here can be adapted to the Gaussian setup, where they can be used to yield sharper versions of CLT convergence. For a discussion of this application, we refer the reader to [2].

As will be seen, the critical value of $k(N)$ and the structure of $\nu^{(N)}$ depend crucially on the behaviour of the function $H(\cdot|\mu) - F(\cdot)$ near its minima. In particular, if the minimum is unique, say μ^*, then $\nu^{(N)} = (\mu^*)^{\otimes N}$, with $k(N) = o(N)$ if this minimum is non-degenerate. Thus, with a non-degenerate unique minimum, one concludes the propagation of chaos for blocks of size $o(N)$.

This study is related to the one in [6], where a similar question in the case of *Gibbs conditioning* was considered. Due to the extra regularity provided by integration over F, the results here are more satisfactory in that they cover (with sharp rates) genuinely infinite dimensional situations.

The structure of this article is as follows. In the rest of this introduction, we describe our results for a (simple) problem, the Curie-Weiss model. This model exhibits already a range of interesting phenomena. We describe the precise assumptions we put on the function F, the statement of our main results in the non-degenerate case, and an application to the Langevin dynamics of interacting particles, in Section 2. Section 3 is devoted to the proofs of the non-degenerate case, while Section 4 is devoted to the statements and proofs for the degenerate case.

Turning now to the Curie-Weiss model, let $S = \{-1, 1\}$, $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$, $F(\nu) = F_{\alpha,\beta}(\nu) = \beta\langle\nu, x\rangle^2 + \alpha\langle\nu, x\rangle$ and

$$\mu^{(N)}(dx_1, \dots, dx_N) = Z_N^{-1} e^{NF_{\alpha,\beta}(L_N^{\mathbf{x}})}\mu^{\otimes N} = Z_N^{-1} e^{\beta N^{-1}\sum_{i,j=1}^N x_i x_j + \alpha \sum_{i=1}^N x_i}\mu^{\otimes N}(dx_1, \dots, x_N).$$

We distinguish between the following cases

$$\begin{array}{rl} \text{I.} & \alpha = 0, \quad \beta < \frac{1}{2} \quad \text{or} \quad \alpha \neq 0 \\ \text{II.} & \alpha = 0, \quad \beta = \frac{1}{2} \\ \text{III.} & \alpha = 0, \quad \beta > \frac{1}{2} \end{array}$$

Let $I(\nu) = H(\nu|\mu) - F_{\alpha,\beta}(\nu)$. A degenerate minima in this context is a minimum μ^* with $I''(\mu^*) = 0$. It is easy to check (by embedding $M_1(S)$ into $\mathbb{R}$) that in both cases I and II, μ^* is unique, with case I corresponding to a non-degenerate μ^* while case II is degenerate. On the other hand, case III corresponds to the case of two non-degenerate minima μ_1^*, μ_2^* (in all cases, and for any minimum μ^*, $\bar{\mu}^* \triangleq \mu^*(1)$ satisfies the relation $\log(\bar{\mu}^*/(1-\bar{\mu}^*)) = 4\beta(2\bar{\mu}^* - 1) + 2\alpha\bar{\mu}^*$, with $\bar{\mu}^* = 1/2$ for $\alpha = 0$ and $\beta \leq 1/2$). A corollary of our general results in this paper is the

Corollary 1.

Case I. *Let $k(N) = o(N)$. Then*

$$H(\mu^{N,k(N)} \,|\, (\mu^*)^{\otimes k(N)}) \underset{N\to\infty}{\longrightarrow} 0\,.$$

Case II. *Let $k(N) = o(N^{\frac{1}{2}})$. Then*

$$H(\mu^{N,k(N)} \,|\, \mu^{\otimes k(N)}) \underset{N\to\infty}{\longrightarrow} 0\,.$$

Case III. *Let $k(N) = o(N)$. Then*

$$\lim_{N\to\infty} H(\mu^{N,k(N)} \,|\, \frac{1}{2}\sum_{j=1}^{2} (\mu_j^*)^{\otimes k(N)}) \underset{N\to\infty}{\longrightarrow} 0\,.$$

The possibility of working with $k = o(N)$ in case III was pointed out to us by A. Dembo, who also provided a proof based on the general results [7].

For a study, from a different viewpoint, of the statistics of large blocks of variables in the critical case, we refer the reader to [10] and the references therein. F. Comets has kindly pointed out to us that the rate $k(N) = o(N^{\frac{1}{2}})$ in the critical case II is optimal. Indeed, fix $1 \geq \epsilon > 0$, and let $k(N) = \epsilon\sqrt{N}$ and $\Theta_N := k(N)^{-1/2}\sum_{i=1}^{k(N)} X_i$. Then, by [10, Theorem 1] (taking there $p = \sqrt{\epsilon}$ and $q = \sqrt{N}\epsilon^{-1}$), there exists a random variable w_N such that the law, under $\mu^{N,k(N)}$, of the random vector $(w_N, \Theta_N - \sqrt{\epsilon} w_N)$ converges weakly to a non-degenerate product measure, with first marginal non-Gaussian and second marginal Gaussian. In particular, Θ_N possesses a non-Gaussian limit law for any $\epsilon > 0$. On the other hand, under the law $\mu^{\otimes k(N)}$ the standard CLT implies that its limit law is Gaussian.

Although Corollary 1 is a direct application of our general Theorems 1–4, we provide a (direct) proof of parts I and II in the end of Section 4. Thus, the reader interested in understanding first this simple situation may skip directly there, avoiding the use of the results in [3] or [4].

Acknowledgement We thank A. Dembo for useful suggestions concerning the multiple minima situation, and an anonymous referee for useful comments which led to a simplification of our arguments.

2 The non-degenerate case.

To state our results, we need to introduce some notations, following [1]. Let $I(\nu) = H(\nu|\mu) - F(\nu)$. Denote $m_o = \inf_{\nu \in M_1(S)} I(\nu)$. Let

$$\mathcal{K}_F = \{\nu \in M_1(S) : \ H(\nu|\mu) - F(\nu) = m_o\} .$$

Because F is bounded and continuous, $\mathcal{K}_F$ is nonempty and compact. We often need the assumption of uniqueness of minimizers, summarized as:

(A1) There exists a *unique* $\mu^* \in \mathcal{K}_F$.

(See however Theorem 2 for a discussion of the case where (A1) does not hold).

Let $E_k(S) = \tilde{\otimes}_\pi^k C_b(S)$, the k-fold (projective) tensor product of $C_b(S)$. Let $E_\infty(S) = \oplus_{k=1}^\infty E_k(S)$, and let $sE_\infty(S)$ denote the symmetric sub-algebra of $E_\infty(S)$. For any $e \in E_\infty(S)$, $\deg(e)$ is the largest integer $k \geq 1$ with nonzero component in the sum defining e.

Let $V \in sE_\infty(S)$, with $\deg(V) = r, r \geq 2$, and components $V_k, k = 2, \ldots r$. Let

$$F(\nu) = \sum_{k=2}^r \langle V_k, \nu^{\otimes k} \rangle .$$

(We do not need to consider linear components in the definition of $F(\cdot)$, for these can be eliminated by a suitable modification of the measure μ).

Note that

$$NF(L_N) = \sum_{k=2}^{r} N^{-(k-1)} \sum_{1 \le i_1 \le \ldots i_k \le N} V_k(x_{i_1}, \ldots, x_{i_k}) \,.$$

Define next, on $L^2_0(S, \mu^*)$ (the space of centered, μ^* square integrable functions), an operator of kernel Ξ such that

$$\langle \Xi f, g \rangle_{L^2_0(S,\mu^*)} = \sum_{k=2}^{r} k(k-1) \langle f \otimes g \otimes 1^{\otimes(k-2)} V_k, (\mu^*)^{\otimes k} \rangle \,.$$

Let $K(\mu^*) = \mathrm{Ker}(\mathrm{Id}_{L^2_0(S,\mu^*)} - \Xi)$. If $K(\mu^*) = \{0\}$, we say that μ^* is non-degenerate. The non-degeneracy condition can be also given a Banach space interpretation, c.f. [1, Lemma 2.19, pg. 96] and (11) below. Note that by [1, Theorem B], there exist at most a finite number of non-degenerate elements in $\mathcal{K}_F$.

Our main result is the following.

Theorem 1. *Assume (A1) and that μ^* is non-degenerate. Then, for some constant C independent of N,*

$$\limsup_{N \to \infty} H\left(\mu^{(N)} \,|\, (\mu^*)^{\otimes N}\right) < C \,. \tag{4}$$

As a consequence, for any $k(N) = o(N)$,

$$H\left(\mu^{N,k(N)} \,|\, (\mu^*)^{\otimes k(N)}\right) \underset{N \to \infty}{\longrightarrow} 0$$

Interesting phenomena occur in the case where $\mathcal{K}_F$ consists of a finite number of non-degenerate minima $\{\mu_1^*, \ldots, \mu_J^*\}$. Let

$$c_i = \frac{\left(\sqrt{\det[Id_{L^2_0(S,\mu_i^*)} - \Xi]}\right)^{-1}}{\sum_{j=1}^{J} \left(\sqrt{\det[Id_{L^2_0(S,\mu_j^*)} - \Xi]}\right)^{-1}} \,.$$

Define next $\mu^{N,*} = \sum_{i=1}^{J} c_i (\mu_i^*)^{\otimes N}$, that is $\mu^{N,*}$ consists of a mixture of product measures with weights c_i. Let $\mu^{k,*}$ denote the restriction of $\mu^{N,*}$ to its first k coordinates, and note that it is also a mixture of product measures. We now have:

Theorem 2. *Assume $\mathcal{K}_F$ consists of a finite number of non-degenerate minima $\{\mu_1^*, \ldots, \mu_J^*\}$. Then, for some constant C which is independent of N,*

$$\limsup_{N \to \infty} H\left(\mu^{(N)} \,|\, \mu^{N,*}\right) < C \,. \tag{5}$$

Further, for any $k = o(N)$,

$$H(\mu^{N,k}|\mu^{k,*}) \underset{N\to\infty}{\longrightarrow} 0\,. \tag{6}$$

Remarks 1. The constants C appearing in Theorem 1 and Theorem 2 can be computed explicitly, and the statements can be strengthened to yield a convergence to C of the relative entropies in (4) and (5).

2. The results presented above extend immediately if $F(L_N)$ is replaced by $g(F(L_N))$, where $g(\cdot) \in C_b^3(\mathbb{R};\mathbb{R})$ and F satisfies the assumptions described in the beginning of this section. In this case, $I(\nu) = H(\nu|\mu) - g(F(\nu))$. This extension allows one to consider interactions based on the empirical mean which are not necessarily polynomial. Technical improvements, in particular on the boundedness assumption on the support of μ, are possible but will not be discussed here.

3. By [1, Corollary 1.6], one may state the assumptions on the potential F directly in terms of the Banach spaces B appearing in the course of the proof of Theorem 1. We chose not to do so as the introduction of the functions V allows for a rather explicit expression for the non degeneracy condition.

4. As mentioned in the introduction, (6) strengthens the weak convergence announced in [1, Theorem B], which by itself precludes the existence of propagation of chaos. Unfortunately, the simple reduction from (5) to (6) used in the case of single minimum does not work in the case of non-product measures, and a slightly more involved argument is needed. The proof given below of (6) is based on a suggestion of A. Dembo.

Finally, we show how to apply Theorem 1 in a dynamic setting to deduce propagation of chaos for a system of interacting particles obeying a Langevin dynamic. Let $X_t^{i,N}$ satisfy the system of SDE's [1]

$$dX_t^{i,N} = dB_t^{i,N} - \nabla U(X_t^{i,N})dt + \frac{1}{N}\sum_{i\le j\le N} \nabla_1 V(X_t^{i,N}, X_t^{j,N})dt\,. \tag{7}$$

Here, $(B^{i,N})_{1\le i\le N}$ denote N dependent, $\mathbb{R}^d$-valued Brownian motions, $U : \mathbb{R}^d \to \mathbb{R}^d$ is a C_b^2 function, and $V \in \tilde{\otimes}_\pi^2 C_b^2(\mathbb{R}^d, \mathbb{R})$, which is assumed to admit the representation

$$V = \int_{\mathcal{C}} g(\cdot,\tau)^{\otimes 2} v(d\tau)\,,$$

with $(\mathcal{C}, v)$ a compact space, and $g : \mathbb{R}^d \times \mathcal{C} \to \mathbb{R}$ differentiable in its first coordinate for all $\tau \in \mathcal{C}$. (Note that by [1, Corollary 1.6], such a representation with continuous g always exists, and the only additional restriction here involves its differentiability. It holds, e.g., if $V(x,y)$ possesses a Fourier transform in $L^1(\mathbb{R}^{2d})$). Finally, we assume

$$\mu_0^{(N)} = \mathrm{Law}(\{X_0^{i,N}\}_{1\le i\le N}) = Z_N^{-1} \exp\left(N^{1-r} \sum_{1\le i_1\le\ldots\le i_r\le N} \overline{S}(x_{i_1},\ldots,x_{i_r})\right) \mu_0^{\otimes N}(dx_1,\ldots,dx_N)\,,$$

with $\overline{S} \in \tilde{\otimes}_\pi^r C_b(\mathbb{R}^d, \mathbb{R})$, $r \geq 2$. Denote by $\mathcal{L}^{k,N} \in M_1(C([0,T];(\mathbb{R}^d)^k))$ the law of $(X^{1,N}_\cdot, \ldots, X^{k,N}_\cdot)$. Next, introduce the McKean-Vlasov nonlinear diffusion: let $\overline{B}_t$ denote an $\mathbb{R}^d$ Brownian motion, independent of $\overline{X}_0$, and for $\mu \in M_1(\mathbb{R}^d)$, let

$$
\begin{aligned}
d\overline{X}_t &= d\overline{B}_t - \nabla U(\overline{X}_t)dt + \int \nabla_1 V(\overline{X}_t, y)\overline{u}_t(dy)dt \\
\overline{u}_t &= \mathrm{Law}(\overline{X}_t), \qquad \overline{u}_0 = \mu\,.
\end{aligned}
$$

Under the above assumptions, the process $\overline{X}_t$ exists and is well defined [8]. Let $\overline{P}_T(\mu) \in M_1(C([0,T], \mathbb{R}^d))$ denote its law. We are now ready to state our main result concerning the dynamics.

Theorem 3. *Let $F(\nu) = \langle \overline{S}, \nu^{\otimes r}\rangle$, and $I(\nu) = H(\nu|\mu_0) - F(\nu)$. Assume that (A1) holds with μ^* non-degenerate. Then, there exists a constant C such that*

$$H(\mathcal{L}^{N,N}|\overline{P}_T^{\otimes N}(\mu^*)) \leq C\,.$$

Further, for $k(N) = o(N)$, propagation of chaos holds, that is

$$H(\mathcal{L}^{k(N),N}|\overline{P}_T^{\otimes k(N)}(\mu^*)) \to_{N\to\infty} 0\,.$$

The following corollary is an immediate consequence of Theorem 3.

Corollary 2. *Under the assumptions of Theorem 3, let $\bar{S} = 0$, that is at $t = 0$ the $X^{i,N}$ are i.i.d.. Then, with $k(N) = o(N)$,*

$$H(\mathcal{L}^{k(N)}|\overline{P}_T^{\otimes k(N)}(\mu^*)) \to_{N\to\infty} 0\,.$$

In particular, at time T independence is still preserved for sub-blocks of size $k(N) = o(N)$.

3 Proofs

Proof of Theorem 1

Our proof relies on the sharp Laplace asymptotics derived in [1], which in turn build on [3]. By definition,

$$
\begin{aligned}
B_N &= H(\mu^{(N)}|(\mu^*)^{\otimes N}) = N\left[\frac{\int F(L_N^{\mathbf{x}})e^{NF(L_N^{\mathbf{x}})}d\mu^{\otimes N} - \int \left\langle L_N^{\mathbf{x}}, \log\frac{d\mu^*}{d\mu}\right\rangle e^{NF(L_N^{\mathbf{x}})}d\mu^{\otimes N}}{Z_N} - \frac{\log Z_N}{N}\right] \\
&= N\left[-\left(m_0 + \frac{\log Z_N}{N}\right) + \frac{\int(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))e^{NF(L_N^{\mathbf{x}})}d\mu^{\otimes N}}{Z_N}\right] \triangleq B_N^{(1)} + B_N^{(2)}, \qquad (8)
\end{aligned}
$$

where $\overline{F}(\nu) = F(\nu) - \left\langle \nu, \log \frac{d\mu^*}{d\mu} \right\rangle$.

The sharp asymptotics for the partition function of a mean field model were computed in [1]. Indeed, [1, Theorem B] yields that

$$|B_N^{(1)}| = N \left| m_0 + \frac{\log Z_N}{N} \right| \underset{N\to\infty}{\longrightarrow} \left| \log \sqrt{\det[Id_{L_0^2(S,\mu^*)} - \Xi]} \right| < \infty \,. \tag{9}$$

Hence, the proof of Theorem 1 follows as soon as we show that $B_N^{(2)}$ is uniformly bounded.

By our assumptions, c.f. [1, Theorem 2.13 and Remark 2.16], $\log \frac{d\mu^*}{d\mu} \in C_b(S)$. Hence, $\overline{F}(\nu) = \sum_{k=1}^r \langle V_k, \nu^{\otimes k} \rangle$, with $V_1 \in C_b(S)$, and $V = (V_1, \ldots, V_r) \in sE_\infty$ with $\deg(V) = r$.

We next follow the procedure suggested in [1] in order to embed the computation of $B_N^{(2)}$ into a Banach space, for which the results of [3] may be applied. By [1], Corollary 1.6, there exists a compact measure space $(\mathcal{C}, v)$, a continuous function $h : \mathcal{C} \to C_b(S)$ such that

$$F(\mu) = \langle \mathcal{V}, \mu^{\otimes r} \rangle$$

with

$$\mathcal{V} = \int_{\mathcal{C}} h^{\otimes r}(\tau) v(d\tau) \,.$$

Clearly, one also has

$$\overline{F}(\mu) = \langle \overline{\mathcal{V}}, \mu^{\otimes r} \rangle$$

for an appropriate $\overline{\mathcal{V}}$ (but same $(\mathcal{C}, v)$). Let $B = L^r(\mathcal{C}, v)$. B is a type 2 Banach space (c.f. [1]). Define $T_h : \; M_1(S) \to B$ by

$$T_h(\mu) : \tau \mapsto \int_S h(\tau)(x) \mu(dx)$$

and $\mathcal{P} : B \to \mathbb{R}$ by

$$\mathcal{P}(\phi) = \int_{\mathcal{C}} \phi^r dv \,.$$

With these definitions, $F = \mathcal{P} \circ T_h$.

By the fact that μ^* is a minimizer (see [1, Theorem 2.13 and Remark 2.16]), it holds that for any neighborhood O of μ^*, there exists a constant $K = K(O)$ such that, whenever $L_N^{\mathbf{x}} \in O$,

$$|N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| \le K |T_h(\sqrt{N}(L_N^{\mathbf{x}} - \mu^*))|_B^2 \,. \tag{10}$$

(See (16) below for an understanding of why (10) holds true). Next, let α denote the law induced by μ on B by the map $x \mapsto h(\cdot)(x)$. Note that $T_h(\sqrt{N}(L_N^{\mathbf{X}} - \mu^*)) = \sqrt{N}(T_h(L_N^{\mathbf{X}} - \mu^*))$, whereas the linearity of T_h as a map on $M_1(S)$ implies that $T_h(L_N^{\mathbf{X}} - \mu^*)$ is an empirical mean of i.i.d.(α), B–valued random variables for which the results of Bolthausen [3] apply. In particular, let (i, H, B) denote the abstract Wiener space generated an B, and denote the Hilbert-norm by H (c.f. [3] for details of the construction). Then (c.f. [1, Lemma 2.19, pg. 96]), the hypothesis of non degeneracy is nothing but the statement that, for $\hat{\varphi} \neq 0$,

$$D^2\mathcal{P}(T_h\mu^*)(\hat{\varphi}, \hat{\varphi}) < |\varphi|_H^2 \tag{11}$$

where $\varphi \in B^*$, the dual space of the Banach space B, and $\hat{\varphi} = \int x\varphi(x)\mu^*(dx) \in B$.

Returning to the evaluation of $B_N^{(2)}$, note that by (9), for some constant C independent of N, and all N large enough,

$$B_N^{(2)} \leq C \int |N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| e^{N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))} d(\mu^*)^{\otimes N} . \tag{12}$$

By Varadhan's lemma, for any neighborhood O of μ^*,

$$\lim_{N\to\infty} \int |N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| e^{N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))} 1_{\{L_N^{\mathbf{x}} \notin O\}} d(\mu^*)^{\otimes N} = 0 . \tag{13}$$

(13) and the fact that the set

$$\{\nu : |T_h(\sqrt{N}(\nu - \mu^*))|_B < \sqrt{N}/c\} = \{\nu : |T_h(\nu - \mu^*)|_B < 1/c\}$$

is an open set imply that for any constant $c > 0$,

$$\begin{aligned} &\limsup_{N\to\infty} \int |N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| e^{N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))} d(\mu^*)^{\otimes N} \\ &\leq \limsup_{N\to\infty} \int |N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| e^{N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))} 1_{|T_h(\sqrt{N}(L_N^{\mathbf{x}} - \mu^*))|_B < \sqrt{N}/c\}}\, d(\mu^*)^{\otimes N} . \end{aligned} \tag{14}$$

Further, by the bound (10), for every neighborhood O of μ^* and constant c there exists a constant $c_1 = c_1(c)$ such that

$$\limsup_{N\to\infty} \int |N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| e^{N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))} 1_{(|T_h(\sqrt{N}(L_N^{\mathbf{x}} - \mu^*))|_B < c\}}\, 1_{\{L_N^{\mathbf{x}} \in O\}} d(\mu^*)^{\otimes N} \leq c_1(c) e^{c_1(c)} \tag{15}$$

(In fact, the precise limit in (15) can be computed by using the CLT of [1], Theorem B, but we do not need it here).

To conclude the proof of Theorem 1, we borrow an argument from [3], pg. 315. First, note that by the argument there, there exists an ϵ small enough such that, denoting $A_\epsilon = \{x \in B : \frac{1}{2}\, D^2\mathcal{P}(T_h\mu^*)(x, x) +$

$\epsilon|x|^2 \geq 1\}$, one has $\delta(\epsilon) := \frac{1}{2}\inf_{x\in A_\epsilon}|x|_H^2 > 1$. Next, denoting $y_N = T_h(\sqrt{N}(L_N^{\mathbf{x}} - \mu^*)) \in B$, the Frechet differentiability of $\mathcal{P}$ implies that

$$N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*)) = N(\mathcal{P}(T_h(L_N^{\mathbf{x}})) - \mathcal{P}(T_h(\mu^*))) = \frac{1}{2}D^2\mathcal{P}(T_h\mu^*)(y_N, y_N) + K(y_N) \tag{16}$$

with

$$\limsup_{|y/\sqrt{N}|\to 0} \frac{K(y)}{|y|^2} = 0\,.$$

Let c be such that, for $|y_N|_B < \sqrt{N}/c$,

$$|N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))| \leq \frac{1}{2}D^2\mathcal{P}(T_h\mu^*)(y_N, y_N) + \epsilon|y_N|_B^2\,.$$

Then, reducing ϵ further and increasing c, if necessary, one has by Bolthausen's extension of Yurinskii's inequality to Banach spaces, c.f. [3, Theorem 3], that, for all $c < t < \sqrt{N}/c$,

$$(\mu^*)^{\otimes N}(y_N \in \sqrt{t}A_\epsilon) \leq \exp(-\delta(\epsilon)t)\,.$$

Hence, for some constant c_2 independent of N, and all N large enough,

$$\begin{aligned}
&\int_{\{c<|y_N|<\sqrt{N}/c\}} \left|N(\overline{F}(L_N^{\mathbf{x}}) - \overline{F}(\mu^*))\right| e^{N(\overline{F}(L_N^{\mathbf{x}})-\overline{F}(\mu^*))} d(\mu^*)^{\otimes N} \\
&\leq e + \int_0^\infty te^t(\mu^*)^{\otimes N}(y_N \in \sqrt{t}A_\epsilon)dt \leq e + \int_0^\infty te^te^{-\delta(\epsilon)t}dt < c_2\,,
\end{aligned} \tag{17}$$

where the last inequality is due to $\delta(\epsilon) > 1$.

Combining (14), (15) and (17), one concludes that $\sup_N B_N^{(2)} < \infty$, yielding (4). The second part now follows from the estimate (3). $\square$

Proof of Theorem 2

Throughout this proof, C denotes a constant whose value may change from line to line but which is always independent of N. Let

$$B_N = H(\mu^{(N)}|\mu^{N,*})\,.$$

By definition,

$$\begin{aligned}
B_N \quad &= \quad N\left(-\frac{\log Z_N}{N} + \sum_{i=1}^J c_i[F(\mu_i^*) - H(\mu_i^*|\mu)]\right) \\
&\quad + NE^{(N)}\left(\sum_{i=1}^J c_i\left(H(\mu_i^*|\mu) - F(\mu_i^*)\right) - \frac{1}{N}\log\left(\frac{d\mu^{N,*}}{d\mu^{\otimes N}}\right) + F(L_N^{\mathbf{X}})\right) \triangleq B_N^{(1)} + B_N^{(2)}\,,
\end{aligned}$$

where throughout $E^{(N)}$ denotes expectations with respect to the measure $\mu^{(N)}$.

Exactly as in the proof of Theorem 1, one has that $B_N^{(1)}$ is uniformly bounded. Let now O_i denote arbitrary open, disjoint neighborhoods of μ_i^*, then by [1, Theorem B] it holds that $\mu^{(N)}(O_i) \to_{N\to\infty} c_i$.

Localizing on $\cup_{i=1}^J O_i$ by the large deviations for $L_N^{\mathbf{X}}$, one finds that

$$E^{(N)} \log\left(\frac{d\mu^{N,*}}{d\mu^{\otimes N}}\right) = \sum_{i=1}^J E^{(N)}\left(1_{\{L_N^{\mathbf{X}}\in O_i\}} \log\left(\frac{d\mu^{N,*}}{d\mu^{\otimes N}}\right)\right) + o(1)\,.$$

Note that

$$\frac{d\mu^{N,*}}{d\mu^{\otimes N}}(\mathbf{x}) = \sum_{i=1}^J c_i \exp\left(N\langle L_N^{\mathbf{x}}, \log\frac{d\mu_i^*}{d\mu}\rangle\right)\,.$$

Since for $O_i = O_i(\delta)$ small enough and $L_N^{\mathbf{x}} \in O_i$ one has that

$$|\langle L_N^{\mathbf{x}}, \log\frac{d\mu_j^*}{d\mu}\rangle - (H(\mu_i^*|\mu) - H(\mu_i^*|\mu_j^*))| \le \delta\,,$$

one sees that for $\delta < \min_{i\neq j} H(\mu_i^*|\mu_j^*)/2$ and all N large enough it holds that

$$|E^{(N)}\left(1_{\{L_N^{\mathbf{X}}\in O_i\}}\log\left(\frac{d\mu^{N,*}}{d\mu^{\otimes N}}\right)\right) - NE^{(N)}\left(1_{\{L_N^{\mathbf{X}}\in O_i\}}\langle L_N^{\mathbf{X}}, \log d\mu_i^*/d\mu\rangle\right)| = 0(1)\,.$$

Hence, with $\bar{F}_i(\nu) = F(\nu) - \langle \nu, \log(d\mu_i^*/d\mu)\rangle$,

$$B_N^{(2)} = \sum_{i=1}^J E^{(N)}\left(1_{\{L_N^{\mathbf{X}}\in O_i\}} N(\bar{F}_i(L_N^{\mathbf{X}}) - \bar{F}_i(\mu_i^*)\right) + 0(1)\,.$$

The proof that $B_N \le C$, i.e. of (5), now proceeds exactly as the proof of boundedness of $B_N^{(2)}$ in Theorem 1.

Next, let $\mu_i^{(N)} = \mu^{(N)}(\cdot|L_N^{\mathbf{X}} \in O_i) \in M_1(S^N)$. One has

$$\begin{aligned}
H(\mu_i^{(N)}|(\mu_i^*)^{\otimes N}) &= \frac{E^{(N)}\left(\mathbf{1}_{\{L_N^{\mathbf{X}}\in O_i\}}\log\frac{d\mu_i^{(N)}}{d(\mu_i^*)^{\otimes N}}\right)}{\mu^{(N)}(L_N^{\mathbf{X}}\in O_i)} \\
&= -\log Z_N + \frac{NE^{(N)}\left(\mathbf{1}_{\{L_N^{\mathbf{X}}\in O_i\}}\left(F(L_N^{\mathbf{X}}) + \frac{1}{N}\log\frac{d\mu^{\otimes N}}{d(\mu_i^*)^{\otimes N}}(\mathbf{X})\right)\right)}{\mu^{(N)}(L_N^{\mathbf{X}}\in O_i)} - \log\mu^{(N)}(L_N^{\mathbf{X}}\in O_i) \\
&= \frac{1}{\mu^{(N)}(L_N^{\mathbf{X}}\in O_i)}\left(0(1) + NE^{(N)}\left(\mathbf{1}_{\{L_N^{\mathbf{X}}\in O_i\}}(\bar{F}_i(L_N^{\mathbf{X}}) - \bar{F}_i(\mu_i^*))\right)\right) = 0(1)\,, \qquad (18)
\end{aligned}$$

where the last two equalities follow from the convergence $\mu^{(N)}(L_N^{\mathbf{X}} \in O_i) \to_{N\to\infty} c_i > 0$ and the same argument leading to the boundedness of $B_N^{(2)}$ in Theorem 1.

Let now

$$\mu_i^{N,k} = \mu^{(N)}(dx_1, \ldots, dx_k | L_N^{\mathbf{X}} \in O_i) \in M_1(S^k)\,.$$

It follows from (3) and (18) that if $k = k(N) = o(N)$ then

$$H(\mu_i^{N,k} | (\mu_i^*)^{\otimes k}) \le Ck/N \to_{N\to\infty} 0\,. \tag{19}$$

Recall that $H(\mu^{N,k}|\mu^{k,*}) \le H(\mu^{(N)}|\mu^{N,*}) \le C$. Hence, $d\mu^{N,k}/d\mu^{k,*}$ exists, and, denoting by $||a-b||_{\mathrm{var}}$ the variation distance between two measures,

$$\begin{aligned}
&||\mu^{N,k} - \mu^{k,*}||_{\mathrm{var}} \\
= \;& ||\sum_{i=1}^{J} \mu_i^{N,k}\mu^{(N)}(L_N^{\mathbf{X}} \in O_i) + \mu^{(N)}(dx_1, \ldots, dx_k | L_N^{\mathbf{X}} \notin \cup_{i=1}^J O_i)\mu^{(N)}(L_N^{\mathbf{X}} \notin \cup_{i=1}^J O_i) - \sum_{i=1}^{J} c_i (\mu_i^*)^{\otimes k}||_{\mathrm{var}} \\
\le \;& \mu^{(N)}(L_N^{\mathbf{X}} \notin \cup_{i=1}^J O_i) + \sum_{i=1}^{J} |c_i - \mu^{(N)}(L_N^{\mathbf{X}} \in O_i)| + \sum_{i=1}^{J} c_i ||\mu_i^{N,k} - (\mu_i^*)^{\otimes k}||_{\mathrm{var}} \to_{N\to\infty} 0\,,
\end{aligned}$$

where the last limit is due to the inequality $||a-b||_{\mathrm{var}} \le \sqrt{2H(a|b)}$ and (19). One concludes that

$$\int \left| \frac{d\mu^{N,k}}{d\mu^{k,*}} - 1 \right| d\mu^{k,*} \to_{N\to\infty} 0\,,$$

and hence

$$\int \frac{d\mu^{N,k}}{d\mu^{k,*}} d\mu_i^{\otimes k} = 1 + \epsilon_{N,k,i}\,, \quad \max_{j=1}^{J} |\epsilon_{N,k,j}| \underset{N\to\infty}{\longrightarrow} 0\,.$$

Next, with $\bar f = \log(d\mu^{N,k}/d\mu^{k,*}) \in C_b(S^k)$,

$$H(\mu^{N,k}|\mu^{k,*}) = \langle \bar f, \mu^{N,k} \rangle\,, \tag{20}$$

whereas, (at least for N/k integer, the general case following by truncation),

$$\begin{aligned}
H(\mu^{(N)}|\mu^{N,*}) &= \sup_{f \in C_b(S^N)} \left(\langle f, \mu^{(N)} \rangle - \log \int e^f d\mu^{N,*} \right) \\
&\ge \sup_{f \in C_b(S^k)} \left(\frac{N}{k} \langle f, \mu^{N,k} \rangle - \log \int e^{(f(x_1,\ldots,x_k)+\ldots+f(x_{N-k+1},\ldots,x_N))} d\mu^{N,*} \right) \\
&\ge \left(\frac{N}{k} \langle \bar f, \mu^{N,k} \rangle - \log \int e^{(\bar f(x_1,\ldots,x_k)+\ldots+\bar f(x_{N-k+1},\ldots,x_N))} d\mu^{N,*} \right)
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{N}{k}\langle \bar{f}, \mu^{N,k}\rangle - \log \sum_{i=1}^{J} c_i \left(\int e^{\bar{f}} d(\mu_i^*)^{\otimes k} \right)^{N/k} \\
&= \frac{N}{k}\langle \bar{f}, \mu^{N,k}\rangle - \log \sum_{i=1}^{J} c_i (1+\epsilon_{N,k,i})^{N/k} \\
&\geq \frac{N}{k}\langle \bar{f}, \mu^{N,k}\rangle - \frac{N}{k} \max_{i=1}^{J} |\epsilon_{N,k,i}| \,. \qquad (21)
\end{aligned}
$$

Combining (20) and (21), one arrives at

$$
H(\mu^{N,k}|\mu^{k,*}) \leq \frac{k}{N} H(\mu^{(N)}|\mu^{N,*}) + \max_{i=1}^{J} |\epsilon_{N,k,i}| \underset{N\to\infty}{\longrightarrow} 0 \,,
$$

as soon as $k = o(N)$. □

Proof of Theorem 3 By [1, Page 115, (3.6)], one sees that the law $\mathcal{L}^{N,N}$ is Gibbsian with respect to $\overline{P}_T^{\otimes N}(\mu_0)$ (with $S = C([0,T], \mathbb{R}^d)$, and bounded, continuous potential $F \in \tilde{\otimes}_\pi^s C_b(S)$, $s = r \vee 3$, see [1, Page 118, Corollary 3.21]). Hence, Theorem 3 follows from Theorem 1 as soon as one establishes that the variational problem associated with F possesses a unique, non degenerate minimum which is $\overline{P}_T(\mu^*)$. But this is the content, respectively, of [1, Page 136, Corollary 3.10] and [1, Page 122, Theorem 2.7]. □

4 The degenerate case

For simplicity, we assume (A1) in the degenerate case. Let $d = \dim K(\mu^*)$, let $\{e_j\}_{j=1}^d$ denote an orthonormal base of $K(\mu^*)$ in $L_0^2(S,\mu^*)$, and define $\lambda(t)$ by

$$
\lambda(t) = \inf\{I(\nu) - m_o \,:\, \langle e_j, \nu\rangle = \langle e_j, \mu^*\rangle + t_j \,,\ 1 \leq j \leq d\} \,,\ \ t \in \mathbb{R}^d \,.
$$

We refer to [1] (notably, the proof of Theorem C) for properties of $\lambda(t)$.

In order to state our last hypothesis, we refer to the construction of the embedding into the space B described in the proof of Theorem 1. Let $\{\phi_j\}, j = 1, \ldots, d$ denote an orthonormal basis in the d-dimensional space of $\phi \in B^*$ achieving equality in (11), such that $\phi_j(T_h(\nu - \mu^*)) = \langle e_j, \nu - \mu^*\rangle$. Defining $q : B \to \mathbb{R}$ by $q(x) = (\phi_1(x), \ldots, \phi_j(x))$, and $\Lambda : B \to \mathbb{R}$ by $\Lambda(\phi) = \inf\{H(\nu|\mu) \,:\, T_h(\nu) = \phi\}$, it holds that (see [4, Pg. 172], and use [1, Lemma 3.8] to identify $\lambda(t)$ of [1] with $\lambda(t)$ of [4])

$$
\lambda(t) = \sup\{F(x + T_h(\mu^*)) - \Lambda(x + T_h(\mu^*)) - m_o \,:\, q(x) = t\} \,.
$$

In order to apply [4] (i.e., to obtain certain local limit theorems uniformly), we will need the following smoothness hypothesis on the finite dimensional measures $\mu \circ q^{-1}$.

(A2) The characteristic function of the measure $\mu \circ q^{-1}$ on $\mathbb{R}^d$ is in L_p, some $\infty > p \geq 1$.

Theorem 4. *Assume (A1) and that μ^* is degenerate. Further, assume (A2) and that $\lambda(t) \sim_{t\to 0} L||t||^p$, some $p \geq 3$ integer and $L > 0$.*

Then,

$$H\left(\mu^{(N)} \,|\, (\mu^*)^{\otimes N}\right) = 0(N^{1-2/p})\,. \tag{22}$$

As a consequence, for any $k(N) = o(N^{2/p})$,

$$H\left(\mu^{N,k(N)} \,|\, (\mu^*)^{\otimes k(N)}\right) \underset{N\to\infty}{\longrightarrow} 0\,.$$

Proof of Theorem 4 .

Throughout this proof, C denotes a constant whose value may change from line to line but which is always independent of N. We essentially follow the proof of Theorem 1, whose notations we adopt, except that one has to condition on the degenerate directions, as in [4]. Here, with B_N defined in (8), it is enough to prove that $B_N = 0(N^{1-2/p})$. By [1], Theorem C, we know that $Z_N e^{-Nm_o} = 0(N^{d(1/2-1/p)})$, and hence $B_N^{(1)} = 0(\log(N))$. Therefore, (12) is replaced by

$$B_N^{(2)} \leq CN^{-d(\frac{1}{2}-\frac{1}{p})} \int N|\bar{F}(L_N^{\mathbf{x}}) - \bar{F}(\mu^*)| e^{N(\bar{F}(L_N^{\mathbf{x}})-\bar{F}(\mu^*))} d(\mu^*)^{\otimes N}\,. \tag{23}$$

By standard large deviations as in (13), one can localize the computation of (23) to any fixed neighborhood O of μ^*.

Letting now $q(T_h(L_N^{\mathbf{x}} - \mu^*)) = t(\mathbf{x})$, one may write $T_h(L_N^{\mathbf{x}} - \mu^*) = V_{t(\mathbf{x})} + W_{t(\mathbf{x})}$ with $q(V_{t(\mathbf{x})}) = t$ and $q(W_{t(\mathbf{x})}) = 0$. Further, since $|N(\bar{F}(L_N^{\mathbf{x}}) - \bar{F}(\mu^*)| \leq C|T_h(\sqrt{N}(L_N^{\mathbf{x}} - \mu^*))|_B^2$, it holds that $|N(\bar{F}(L_N^{\mathbf{x}}) - \bar{F}(\mu^*)| \leq CN|V_{t(\mathbf{x})}|^2 + CN|W_{t(\mathbf{x})}|_B^2$. Fix now $c > 0$ large enough such that [4, Proposition 3.12] can be applied, and take the neighborhood O to be such that $L_N^{\mathbf{x}} \in O \Rightarrow |V_{t(\mathbf{x})}|_B^2 + |W_{t(\mathbf{x})}|_B^2 < c^{-1}$. We now have

$$\int N|W_{t(\mathbf{x})}|_B^2 e^{N(\bar{F}(L_N^{\mathbf{x}})-\bar{F}(\mu^*))} \mathbf{1}_{\{L_N^{\mathbf{x}} \in O\}} d(\mu^*)^{\otimes N} \tag{24}$$

$$\leq cZ_N e^{-Nm_o} + \int_{\{c \leq N|W_{t(\mathbf{x})}|_B^2 \leq N/c\}} N|W_{t(\mathbf{x})}|_B^2 e^{N(\bar{F}(L_N^{\mathbf{x}})-\bar{F}(\mu^*))} \mathbf{1}_{\{L_N^{\mathbf{x}} \in O\}} d(\mu^*)^{\otimes N} \leq CZ_N e^{-Nm_o}\,,$$

where the last inequality is due to the uniform CLT contained in [4, Proposition 3.12] (note that we are working in the non-degenerate directions, and follow the same computation as in Theorem 1 when deriving

(17)). On the other hand, following the computation leading to [4, (4.2)–(4.6)], one finds that

$$\int N|V_{t(\mathbf{x})}|_B^2 e^{N(\bar{F}(L_N^{\mathbf{x}})-\bar{F}(\mu^*))} d(\mu^*)^{\otimes N} \le CN^{d/2} \int (N|t|^2) e^{-LN|t|^p} dt \le CN^{d(\frac{1}{2}-\frac{1}{p})+\frac{p-2}{p}} .$$

Combining the last inequality with (24) and (23) yields the desired estimate on $B_N^{(2)}$ and hence the theorem.

□

Proof of Corollary 1

We consider the case $\alpha = 0$, the general case being similar. The argument leading from Case III to case I being exactly as outlined in the proof of Theorem 2, we deal with cases I and II only. Throughout this proof, C denotes a constant whose value may change from line to line but which is independent of N. Recall that

$$B_N = N \left[\frac{\int F(L_N^{\mathbf{x}}) e^{NF(L_N^{\mathbf{x}})} d\mu^{\otimes N}}{Z_N} - \frac{\log Z_N}{N} \right] .$$

By theorem 2 of [9], one knows that $Z_N = C(1+o(1))$ (case I) while $Z_N = CN^{1/4}(1+o(1))$ (case II) for some constant C (which may be computed, although we do not need this computation). Let $Y_N = \sqrt{N}\langle L_N, x\rangle$. Then

$$B_N = O(\log Z_N) + \beta Z_N^{-1} \int Y_N^2 e^{\beta Y_N^2} d\mu^{\otimes N} .$$

Next, by Chebycheff's inequality, denoting $\Lambda(\theta) = \log E(e^{\theta X_1}) = \log\cosh\theta$, one has

$$E(Y_N^2 e^{\beta Y_N^2}) \le 2\int_0^{\sqrt{N}} z^2 e^{\beta z^2} e^{\left(N\Lambda\left(\frac{z}{\sqrt{N}}\right)-z^2\right)} dz \le 2\int_0^{\infty} z^2 e^{\beta z^2} e^{\left(N\Lambda\left(\frac{z}{\sqrt{N}}\right)-z^2\right)} dz .$$

A direct computation reveals that $\Lambda'(0) = \Lambda^{(3)}(0) = 0$, while $\Lambda''(0) = 1$, and

$$\gamma = \frac{\max_{0\le\xi\le 1} \Lambda^{(5)}(\xi)}{5!} + \frac{\Lambda^{(4)}(0)}{4!} < 0 .$$

Hence, in Case I, one gets

$$E(Y_N^2 e^{\beta Y_N^2}) \le 2\int_0^{\infty} z^2 e^{-(\frac{1}{2}-\beta)z^2} dz \le C$$

while in Case II one gets

$$E(Y_N^2 e^{Y_N^2/2}) \le 2\int_0^{\infty} z^2 e^{-\gamma z^4/N} dz \le 2N^{3/4} \int_0^{\infty} z^2 e^{-\gamma z^4} dz = CN^{3/4} .$$

One conclude that $B_N = 0(1)$ (Case I) and $B_N = 0(N^{1/2})$ (Case II). The conclusion of the theorem follows.

□

References

[1] G. Ben Arous and M. Brunaud, *Methode de Laplace: etude variationelle des fluctuation de diffusions de type "champ moyen"*, Stochastics 31 (1990), pp. 79–144.

[2] G. Ben Arous, Lecture notes, St. Flour (1995).

[3] E. Bolthausen, *Laplace approximations for sums of independent random vectors*, PTRF 72 (1986), pp. 305–318.

[4] E. Bolthausen, *Laplace approximations for sums of independent random vectors, part II: degenerate maxima and manifold of maxima*, Prob. Theory Rel. Fields 76 (1987), pp. 167–206.

[5] I. Csiszár, *Sanov property, generalized I-projection and a conditional limit theorem*, Ann. Probab. 12 (1984), pp. 768–793.

[6] A. Dembo and O. Zeitouni, *Refinements of the Gibbs conditioning principle*, Prob. Theory Rel. Fields 104 (1996), pp. 1–14.

[7] P. Diaconis and D. Freedman, *Finite exchangeable sequences*, Ann. Probab. 8 (1980), pp. 745–764.

[8] T. Funaki, *A certain class of diffusion processes associated with nonlinear parabolic equations*, Z. Wahr. verw. Geb. 67 (1984), 331–348.

[9] A. Martin-Löf, *A Laplace approximation for sums of independent random variables*, Z. Wahr. ver. Gebiete 59 (1982), pp. 101–115.

[10] F. Papangelou, *On the Gaussian fluctuations of the critical Curie-Weiss model in statistical mechanics*, Prob. Theory Rel. Fields 83 (1989), pp. 265–278.