Information Estimates and Markov Random Fields

Francis Comets* Ofer Zeitouni†

October 13, 1998. Revised February 12, 1999.

Abstract

Consider a random field, with values in some finite set $\Sigma \subset \mathbb{R}$ and index set a cube $\Lambda_n \subset \mathbb{Z}^d$. We show that in the vicinity (in the information-theoretic sense) of strongly mixing Markov fields, considering sub-blocks of variables indexed by $\Lambda_m \subset \Lambda_n$, the distribution on this smaller cube can be described precisely, even when the size of the cube grows with n. The general results are then applied to mean field perturbations of Gibbs measures (in particular, mean field perturbations of Ising models). The proofs use entropy arguments as well as (known) result on complete analyticity and mixing for the Ising model.

AMS 1991 subject classification: 60G60, 60F10, 82B20, 82B44.

Key words and phrases: Information, relative entropy, Gibbs fields, Markov fields, weak-mixing, complete analyticity, mean-field

Short Title: Information in Markov fields

Introduction

Information - or relative entropy - indicates how far apart are two probability measures. Though it is not a distance it shares some important properties with Euclidean and Hilbertian norm [3]. Involving a logarithmic density it has a simple expression for Gibbs distributions, and it has natural applications in this context. Given a bound on the information between two random fields on some volume, a natural question is to derive quantitative estimates on the distance between their restrictions to smaller

*Université Paris 7 — Denis Diderot, Mathématiques, case 7012, 2 place Jussieu, 752151 Paris Cedex 05, France. F.C is from UPRES-A CNRS 7055 “Statistique et Modèles Aléatoires”; This work was started when he visited the Technion.

†Department of Electrical Engineering, Technion, Haifa 32000, Israel. Part of this work was done while this author visited the Centre Borel in the IHP.

volumes. This question was addressed in [4], in the context of extending the validity of Gibbs equivalence of ensembles, for the case of a product reference measure, and in [1] for the study of mean-field perturbations of such a measure. Information with respect to a product measure has a nice subadditivity property due to the decomposition formula, but it remains well-behaved in the Markovian case, see for instance [11] where this decomposition is thoroughly exploited and [13] for its counterparts to specific entropy and to the Gibbs variational principle. We also note that [17] discuss a general framework for proving the equivalence of ensembles, based on entropy-like functionals. Finally, in the one dimensional Markov case, [26] and [6] provide conditional limit theorems in the context of the Gibbs conditioning principle and its extensions.

In this paper we study information with respect to a Markov field ν on (large) cubes $\Lambda_n = [-n,n]^d$. Our basic result is an explicit upper estimate of the variational norm $\|\mu-\nu\|_{i+\Lambda_m}$ on translates of smaller cubes Λ_m in terms of the information of μ on Λ_n, under the assumption that ν is weak-mixing (Definition 1 below). Introduced by Dobrushin and Shlosman [9], [10], the weak-mixing property as well as the companion properties of strong-mixing and complete analyticity (Definition 2 below), turned out to be connected with the strong uniqueness of Gibbs fields and their static features ([10]), as well as with spectral properties of their dynamics, i.e. existence of spectral gap and logarithmic Sobolev inequalities (see [20] for a review). In particular these mixing conditions are much less stringent than the Dobrushin condition [14] for uniqueness. For our purpose complete analyticity will be most useful to estimate the information.

A large part of the paper is devoted to applications to mean-field perturbations of Gibbs fields. A typical example is the probability measure on Λ_n (see Example 1 below),

$$d\mu_n = Z_n^{-1}\exp\left(\frac{t|\Lambda_n|}{2}\left(\frac{\sum_{i\in\Lambda_n} x_i}{|\Lambda_n|}\right)^2\right)\, d\nu_n\,,$$

where ν_n is the binary nearest neighbor Ising model with periodic boundary conditions at inverse temperature β without external field, and where Z_n is the normalizing constant. Note that this model is analogous to the Curie-Weiss model, and is a finite volume version of the Kac's asymptotics justifying Van der Waals theory ([16],[2], with $\gamma=\gamma(n)$). In dimension $d=2$ we obtain the following:

1. for $\beta<\beta_c$ the critical value, $t\sigma^2<1$ with σ^2 the variance under $\nu=\lim_n \nu_n$, and $m=m(n)=o(n)$,

$$\lim_{n\to\infty}\|\mu_n-\nu_n\|_{\Lambda_m}=0\quad,$$

2. for $\beta < \beta_c, t\sigma^2 = 1$ and $m = o(n^{1/2})$, the same conclusion holds,

3. for $\beta < \beta_c, t\sigma^2 > 1$ or $\beta \geq \beta_c$, and $m = o(n)$,

$$\lim_{n\to\infty} \|\mu_n - \frac{1}{2}(\nu_n^{h_n} + \nu_n^{-h_n})\|_{\Lambda_m} = 0 \ ,$$

   with $\nu_n^{h_n}$ the periodic boundary condition Ising measure with some appropriate external field h_n.

The following Section 1 is devoted to the general estimate relating information and variational norm. Then, in Section 2, analyticity aspects of Gibbs fields are developed to control entropies. The last section studies the example described above, and concludes with a discussion of some other recent alternative approaches.

1 Main Estimate

Notations: the state space is, for simplicity, a finite set $\Sigma \subset \mathbb{R}$. The results in Section 1 are valid on an abstract polish space Σ, but we will need this restriction on Σ for our examples in the other sections. $V \subset\subset W$ stands for "V is a finite subset of W".

For a random field μ and $\Lambda, V \subset \mathbb{Z}^d$ we denote by μ^Λ a regular version of μ given $(x_i, i \in \Lambda)$, μ_V the restriction μ to V, and $\mu_V^\Lambda = (\mu^\Lambda)_V$. Throughout, we use the same symbol $\otimes$ to denote the usual tensor product, and the skew product between a measure μ_V and a conditional measure $\nu_{V'}^V$, that is the surgery

$$\left(\mu_V \otimes \nu_{V'}^V\right) \ (dx(V), dx(V')) = \nu_{V'}^V \left(dx(V')|x(V)\right) \ \mu_V \left(dx(V)\right) \ . \tag{1.1}$$

Let μ, ν be two probability measures on Σ^Λ, $\Sigma^{\Lambda'}$ respectively, for some $\Lambda, \Lambda' \subset \mathbb{Z}^d$. The *relative entropy* (or information gain) of μ with respect to ν on a subset $V \subset \Lambda \cap \Lambda'$ is defined as

$$H_V(\mu|\nu) = \begin{cases} \int_{\Sigma^V} \log \frac{d\mu_V}{d\nu_V} d\mu_V \, , & \text{if } \mu_V << \nu_V \\ +\infty \, , & \text{otherwise} \ . \end{cases} \tag{1.2}$$

Then $H_V(\mu|\nu) \in [0, -\log\min\{\nu_V(x); x \in \Sigma^V\}] \subset [0, +\infty]$, and the larger its value the more μ, ν differ on V. Though H does not define a distance it has some nice elementary properties, which we group in the lemma below. We refer in the statement of each property to a proof in the literature. Recall that for any $\Lambda \subset \Sigma^{\mathbb{Z}^d}$, the total variation norm on Λ of any signed measure η is defined as

$$\|\eta\|_\Lambda = \sup\{\int f d\eta; f : \Sigma^\Lambda \mapsto \mathbb{R} \text{ measurable}, \|f\|_\infty = 1\} \, , \tag{1.3}$$

and note that $\|\mu - \nu\|_\Lambda \leq 2$ if μ, ν are probability measures.

Lemma 1 *1. For $\Lambda = V \bigcup V', V \bigcap V' = \emptyset$ we have (see [5, Theorem D.13] or [11, Theorem C.3.1])*

$$H_\Lambda(\mu|\nu) = H_V(\mu|\nu) + \mathbf{E}^\mu H_{V'}(\mu^V|\nu^V) = H_V(\mu|\nu) + H_\Lambda(\mu|\mu_V \otimes \nu^V_{V'})\,. \quad (1.4)$$

In particular,

$$H_\Lambda(\mu|\nu) \geq H_V(\mu|\nu)\,. \quad (1.5)$$

2. For any disjoint V_i, and any probability measures ν^i on Σ^{V_i}, $i = 1, 2, \ldots, k$, with $V = \cup_{i=1}^k V_i$, (see [3])

$$H_V(\mu| \otimes_{i=1}^k \nu^i_{V_i}) = H_V(\mu| \otimes_{i=1}^k \mu_{V_i}) + \sum_{i=1}^k H_{V_i}(\mu|\nu^i) \geq \sum_{i=1}^k H_{V_i}(\mu|\nu^i) \quad (1.6)$$

3. Entropy dominates the variation norm (Pinsker inequality, see e.g. [5, Exercise 6.2.17] or [11, Theorem C.3.1])

$$\|\mu - \nu\|^2_\Lambda \leq 2H_\Lambda(\mu|\nu) \quad (1.7)$$

Let $|.|$ denote the l_1-norm on $\mathbb{Z}^d$, and fix some $r \in \mathbb{R}^+$. Let $\overline{V} = \{i \in \mathbb{Z}^d; \exists j \in V : |i-j| \leq r\}$ and $\partial V = \overline{V} \setminus V$. For $n \geq 0$ denote by Λ_n the cube $[-n, n]^d \in \mathbb{Z}^d$.

In what follows ν denotes an *r-Markov random field* on $\Sigma^{\mathbb{Z}^d}$, i.e., there exists a family π of transition kernels - called *specification* - π_V from $\Sigma^{\partial V}$ to Σ^V, $V \subset\subset \mathbb{Z}^d$, such that

$$\nu_V^{V^c} = \pi_V \quad , \quad \nu - a.s. \quad (1.8)$$

Chop $\mathbb{Z}^d$ into translates $\Lambda_{k,j} = j + \Lambda_k$ of Λ_k separated by corridors of width r. Denote by $T_{n,k}$ the set of indices j such that $\overline{\Lambda_{k,j}} \subset \Lambda_n$, and $C_{n,k}$ the complement of $\bigcup_{j \in T_{n,k}} \Lambda_{k,j}$ in Λ_n. Then $(C_{n,k}; \Lambda_{k,j}, j \in T_{n,k})$ is a partition of Λ_n. Moreover from the Markov property (1.8) it follows that

$$\nu_{\Lambda_n} = \left(\otimes_{j \in T_{n,k}} \pi_{\Lambda_{k,j}}\right) \otimes \nu_{C_{n,k}} \quad (1.9)$$

Combining this with (1.4 and 1.6) we obtain

$$\begin{aligned} H_{\Lambda_n}(\mu|\nu) &= \mathbf{E}^{\mu_{C_{n,k}}} H_{\Lambda_n \setminus C_{n,k}} \left(\mu^{C_{n,k}}| \otimes_{j \in T_{n,k}} \pi_{\Lambda_{k,j}}\right) + H_{C_{n,k}}(\mu|\nu) \\ &\geq \sum_{j \in T_{n,k}} \mathbf{E}^{\mu_{C_{n,k}}} H_{\Lambda_{k,j}} \left(\mu^{C_{n,k}}|\pi_{\Lambda_{k,j}}\right) + H_{C_{n,k}}(\mu|\nu) \end{aligned}$$

Using again (1.4) we have:

Proposition 1 *For ν an r-Markov field,*

$$\begin{aligned} H_{\Lambda_n}(\mu|\nu) &\geq \sum_{j \in T_{n,k}} H_{\overline{\Lambda_{k,j}}}\left(\mu|\mu_{\Lambda^c_{k,j}} \otimes \pi_{\Lambda_{k,j}}\right) + H_{C_{n,k}}(\mu|\nu) \\ &\geq \sum_{j \in T_{n,k}} H_{\Lambda_{k,j}}\left(\mu|\mu_{\Lambda^c_{k,j}} \otimes \pi_{\Lambda_{k,j}}\right) + H_{C_{n,k}}(\mu|\nu) \end{aligned} \tag{1.10}$$

These two summands in the right hand side of (1.10) are quite different in nature. The first one contains volume information, and is of order n^d if μ has conditional distribution different from π. On the other hand if the family of specification π exhibits multiplicity of phases, then this term is equal to zero for some Gibbsian $\mu \neq \nu$; hence the second term contains all the discrepancy between the two measures, and is typically of surface order n^{d-1}. In what follows we concentrate on taking advantage of the first term alone. For this purpose a strong uniqueness condition is required for the Markov field [9], [19].

Definition 1 *The Markov specification π is weak-mixing if there exist constants $\gamma > 0$, $C < \infty$ such that for all finite cubes $V \subset \Lambda \subset \mathbb{Z}^d$,*

$$\sup\{\|\pi_\Lambda(.|x) - \pi_\Lambda(.|x')\|_V ; x, x' \in \Sigma^{\Lambda^c}\} \leq C \sum_{i \in V, j \in \partial\Lambda} \exp(-\gamma|i-j|)$$

Theorem 1 *Assume that ν is an r-Markov random field for a weak-mixing specification π. For $1 \leq m \leq n/2$ denote by $A(m,n) = \{j : j + \Lambda_m \subset \Lambda_n\}$ and by $|A(m,n)|$ its cardinality.*
Then there exist $\varepsilon_1(m) = O(1/m)$, $\varepsilon_2(n) = O(\log n/n)$, and $\alpha(m,n)$ with $\alpha(m,n) \sim 2(m/n)^d$ as $m, n \to \infty$, such that

$$|A(m,n)|^{-1} \sum_{j \in A(m+[(2d/\gamma)\log m],n)} \|\mu - \nu\|^2_{j+\Lambda_m} \leq \alpha(m,n) H_{\Lambda_n}(\mu|\nu) + \varepsilon_1(m) \tag{1.11}$$

and

$$|A(m,n)|^{-1} \sum_{j \in A(m,n)} \|\mu - \nu\|^2_{j+\Lambda_m} \leq \alpha(m,n) H_{\Lambda_n}(\mu|\nu) + \varepsilon_1(m) + \varepsilon_2(n) \tag{1.12}$$

for all probability measure μ on Σ^{Λ_n}.

Proof: Let $k = k(m) = m + [(2d/\gamma)\log m]$. Using the notations of Proposition 1, we will consider smaller cubes $\Lambda_{m,j} = j + \Lambda_m$ included in $\Lambda_{k,j}$. We have from (1.5, 1.7, 1.10) that

$$\begin{aligned} H_{\Lambda_n}(\mu|\nu) &\geq \sum_{j \in T_{n,k}} H_{\Lambda_{m,j}}\left(\mu|\mu_{\Lambda^c_{k,j}} \otimes \pi_{\Lambda_{k,j}}\right) \\ &\geq \sum_{j \in T_{n,k}} \|\mu - \mu_{\Lambda^c_{k,j}} \otimes \pi_{\Lambda_{k,j}}\|^2_{\Lambda_{m,j}}/2 \\ &\geq \sum_{j \in T_{n,k}} \left(\|\mu - \nu\|^2_{\Lambda_{m,j}} - 4\|\nu - \mu_{\Lambda^c_{k,j}} \otimes \pi_{\Lambda_{k,j}}\|_{\Lambda_{m,j}}\right)/2 \end{aligned}$$

where the last term can be estimated using the identity $\nu_{\overline{\Lambda_{k,j}}} = \nu_{\partial\Lambda_{k,j}} \otimes \pi_{\Lambda_{k,j}}$ following from the r-Markov assumption, together with the weak-mixing property,

$$\begin{aligned} \|\nu - \mu \otimes \pi_{\Lambda_{k,j}}\|_{\Lambda_{m,j}} &\leq C|\Lambda_m|\,|\partial\Lambda_k| \exp(-\gamma(k-m)) \\ &\leq 2Cdr(2m+1)^d(2k(m)+1)^{d-1}\exp(-\gamma[(2d/\gamma)\log m]) \\ &=: \varepsilon_1(m)/2 \end{aligned}$$

hence $\varepsilon_1(m) = O(m^{-1})$. Shifting the origin we clearly get similar estimates. Adding them up we find

$$(2(k+r)+1)^d H_{\Lambda_n}(\mu|\nu) \geq \sum_{j\in A(k+r,n)} \left(\|\mu-\nu\|^2_{\Lambda_{m,j}}/2 - \varepsilon_1(m)\right) \tag{1.13}$$

and then adding the terms with $\Lambda_{m,j}$ close to the boundary $\partial\Lambda_n$ we finally obtain

$$(2(k+r)+1)^d H_{\Lambda_n}(\mu|\nu) + 2\times 2d(2n+1)^{d-1}(\frac{2d}{\gamma}\log m + r) \geq \sum_{j\in A(m,n)} \left(\frac{\|\mu-\nu\|^2_{\Lambda_{m,j}}}{2} - \varepsilon_1(m)\right).$$

The two estimates above imply the theorem with $\varepsilon_2(n) = O(\log n/n)$, since $m \leq n/2$ and $|A(m,n)| \geq n^d$. □

We have the straightforward

Corollary 1 *Let ν_n be a sequence of r-Markov random fields, all with the same weak-mixing specification π in Λ_n. Let μ_n be a sequence of probability measures on Σ^{Λ_n} such that $H_{\Lambda_n}(\mu_n|\nu_n) = O(1)$. Then for any sequence $m = m(n) = o(n)$, $\|\mu_n - \nu_n\|_{\Lambda_{m,j}} \to 0$ as $n \to \infty$ in density, i.e.,*

$$|A(m,n)|^{-1}|\{j \in A(m,n) : \|\mu_n - \nu_n\|_{\Lambda_{m,j}} > \delta\}| \to 0$$

for all positive δ.

We focus now on translation invariance. For simplicity we assume below that the specification is given by a translation invariant Gibbs potential of finite range r (see definition (2.2) below). Then the finite volume Gibbs measure ν_n on Λ_n with periodic boundary conditions [14] is invariant under the translations of the discrete torus $(\mathbb{Z}/(2n+1)\mathbb{Z})^d$. We consider also translation invariance on $\mathbb{Z}^d$ in the following

Corollary 2 *Assume that π is a Gibbsian specification with translation invariant, finite range potential, and that π is weak-mixing. Denote by ν_n the periodic boundary conditions Gibbs measure on Λ_n, and by ν the unique infinite-volume Gibbs measure. Let μ_n be another probability measure on Σ^{Λ_n} [respectively, $\Sigma^{\mathbb{Z}^d}$] invariant under* *the translations on the discrete torus [respectively, on $\mathbb{Z}^d$]. If $H_{\Lambda_n}(\mu_n|\nu_n) = O(1)$ [respectively, $H_{\Lambda_n}(\mu_n|\nu) = O(1)$] then*

$$\|\mu_n-\nu_n\|_{\Lambda_m} = O\left((\frac{m}{n})^{d/2}\right)+O\left(\frac{1}{m}\right) \qquad \textit{and} \qquad \|\mu_n-\nu\|_{\Lambda_m} = O\left((\frac{m}{n})^{d/2}\right)+O\left(\frac{1}{m}\right)$$

[respectively, $\|\mu_n - \nu\|_{\Lambda_m} = O\left((\frac{m}{n})^{d/2}\right) + O\left(\frac{1}{m}\right)$] for all sequence $m = m(n) = o(n)$.

Proof: Due to translation invariance the summands in the left hand side of (1.11) are equal, and the first statements in the two cases is immediate. To derive the second statement in the torus case it suffices to note that $\|\nu_n - \nu\|_{\Lambda_m}$ tends to zero from the weak-mixing assumption. □

2 Gibbs fields and their analytic properties

A finite range, translation invariant potential is a family $U = \{U_A; A \subset\subset \mathbb{Z}^d\}$ of (bounded) functions from $\Sigma^{\mathbb{Z}^d}$ to $\mathbb{R}$ such that

$$U_A = 0 \text{ if } \mathrm{diam} A > r \ , \ U_A \text{ is } \mathcal{B}_A - \text{measurable}$$

$$U_A = U_{i+A} \circ \theta_i \quad i \in \mathbb{Z}^d, A \subset\subset \mathbb{Z}^d \tag{2.1}$$

with θ_i the shift of vector i on the configuration space. The set of such potentials is a finite dimensional space with the norm $\|U\| = \sup_{A,x} |U_A(x)|$. We assume from now on that the specification π is *Gibbsian* with potential U, i.e.

$$\pi_V\left(dx(V)|x(V^c)\right) = Z_V(U|x(V^c))^{-1} \ \exp\{\sum_{A:A\bigcap V\neq\emptyset} U_A(x)\} \ \alpha^{\otimes V}(dx(V))\,,$$

$$Z_V(U|x(V^c)) = \int_{\Sigma^V} \exp\{\sum_{A:A\bigcap V\neq\emptyset} U_A(x)\} \ \alpha^{\otimes V}(dx(V)) \tag{2.2}$$

for some probability measure α on Σ which is not a point mass. We have denoted by x the configuration which coincides with $x(V)$ [respectively $x(V^c)$] on V [respectively V^c]. Note that conversely the kernels π given by (2.2) define a Markov specification, and a solution ν to (1.8) is called infinite-volume Gibbs measure. Finally, we denote by U^h the potential $U + hX$, $h \in \mathbb{R}$, where $X_i(x) = x_i$ (recall that $\Sigma \subset \mathbb{R}$), $X_A = 0$ if $|A| \neq 1$.

Definition 2 *1. For a potential U we consider the following property: $\exists\varepsilon > 0$, $C < \infty$ such that*

$$|\log Z_V(U + \tilde{U}|x(V^c))| \leq C|V|\,, \tag{2.3}$$

for all complex valued potential $\tilde{U}$ *with range* r *and* $\|\tilde{U}\| < \varepsilon$, *all (finite) cube* V *in* $\mathbb{Z}^d$ *and all boundary conditions* x *on* V^c. *Define the main component as the largest open set in the space of potentials which contains the origin and where the property (2.3) holds. A potential* U *(or the specification* π*) is* completely analytic *if* U *belongs to the main component.*

2. *The Gibbs specification defined by the potential* U *is* completely analytic for the mean *if (2.3) holds merely for the potential* zX, *for all* z *complex such that* $|z| < \epsilon$, *that is*

$$|\log Z_V(U + zX|x(V^c))| \leq C|V|, \tag{2.4}$$

for all (finite) cube V *in* $\mathbb{Z}^d$ *and all boundary conditions* $x(V^c)$.

Property (2.3), sometimes called "restricted complete analyticity" for it is restricted to cubes V, implies uniqueness of the solution ν of (1.8) [9]. Note that $\tilde{U}$ is not assumed translation invariant (recall (2.1)). In view of the shift invariance of the potential U, this unique solution ν is translation invariant. Moreover this complete analyticity property implies the weak mixing property of Definition 1 in the case of a finite state space Σ, see [10]. Note also that (2.4) corresponds to the complete analyticity property for complex perturbations in the magnetic field only.

The example we develop below is a mean-field perturbation of a Markov random field, see (3.1). For the sake of simplicity we have considered Gibbs fields with finitely many real values, and the simplest non-trivial perturbation, a quadratic function of the mean

$$\bar{x}_n = |\Lambda_n|^{-1} \sum_{i \in \Lambda_n} x_i$$

It is well known that the law of the empirical mean $\bar{x}_n$, $\nu(\bar{x}_n \in \cdot)$, satisfies a large deviation principle with (good) rate function $I(y) = \sup_{t \in \mathbb{R}} \{ty - \Gamma(t)\}$ where

$$\Gamma(t) = \lim_{n \to \infty} \Gamma_n(t) \quad , \quad \Gamma_n(t) = |\Lambda_n|^{-1} \log \mathbf{E}^\nu \exp\{t|\Lambda_n|\bar{x}_n\} \quad , \quad t \in \mathbb{R} \tag{2.5}$$

see e.g. the Notes about Chapter IV in [7]. In fact I, Γ depend on π but not on the particular Gibbs measure ν compatible with π. In order to compute informations we need the technical estimates in the next two Propositions.

Proposition 2 *Assume that the Gibbs specification defined by* U *is completely analytic for the mean. Then,*

1. Γ *is analytic on* $(-\varepsilon, \varepsilon)$, I *is analytic in a neighborhood of* $I^{-1}(0)$.

2. *Let* $\sigma^2 = \Gamma''(0)$. *For all real* t, $t\sigma^2 < 1$ *there exists* $\delta > 0$ *such that, as* $n \to \infty$,

$$\mathbf{E}^\nu \left(\exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n - \mathbf{E}^\nu \bar{x}_n)^2\} \mathbf{1}_{|\bar{x}_n - \mathbf{E}^\nu \bar{x}_n| < \delta} \right) \to (1 - \sigma^2 t)^{-1/2},$$

and

$$\mathbf{E}^\nu\left(|\Lambda_n|(\bar{x}_n-\mathbf{E}^\nu\bar{x}_n)^2\ \exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n-\mathbf{E}^\nu\bar{x}_n)^2\}\mathbf{1}_{|\bar{x}_n-\mathbf{E}^\nu\bar{x}_n|<\delta}\right)\to(1-\sigma^2t)^{-3/2}\sigma^2\,. \tag{2.6}$$

3. *Let $\sigma^2=\Gamma''(0)$ and fix $t=1/\sigma^2$. With $\sigma_n^2=\Gamma_n''(0)$, and $t_n=1/\sigma_n^2$, it holds that $t_n\to t$. Assume further that U is completely analytic for the mean, that $\Gamma_n^{(3)}(0)=0,\Gamma^{(4)}(0)<0$, and that $I^{-1}(0)$ is the unique minimizer of the function*

$$I_t(y):=I(y)-\frac{t}{2}(y-I^{-1}(0))^2\quad .$$

Then there exist constants $C_1,C_2,C_3\in(0,\infty)$ such that

$$C_1|\Lambda_n|^{1/4}\le\mathbf{E}^\nu\left(\exp\{\frac{t_n|\Lambda_n|}{2}(\bar{x}_n-\mathbf{E}^\nu\bar{x}_n)^2\}\right)\le C_2|\Lambda_n|^{1/4}\,, \tag{2.7}$$

and

$$\mathbf{E}^\nu\left(|\Lambda_n|(\bar{x}_n-\mathbf{E}^\nu\bar{x}_n)^2\ \exp\{\frac{t_n|\Lambda_n|}{2}(\bar{x}_n-\mathbf{E}^\nu\bar{x}_n)^2\}\right)\le C_3|\Lambda_n|^{3/4}\,. \tag{2.8}$$

4. *In dimension $d=2$, the estimates (2.7, 2.8) hold with t replacing t_n provided that U is completely analytic, that $\Gamma^{(4)}(0)<0$, and that $I^{-1}(0)$ is the unique minimizer of I_t.*

Remark 1 *In the case $\sigma^2t=1$, the assumptions on the derivatives of Γ guarantee that $I^{-1}(0)$ continues -like in the case $\sigma^2t<1$ - to be a local minimizer of the function $I_t(y)$, with a quartic expansion. Computing precise equivalents in (2.7) and (2.8) would require much extra work. Since the constants C are irrelevant for our purpose we give a rough but short proof of (2.8) below.*

In the case $\sigma^2t>1$, $I^{-1}(0)$ is not any more a minimizer of $I_t(\cdot)$, and $\mathbf{E}^\nu x_n$ is not any more the proper centering. To handle this situation, see Proposition 3.

Proof: 1. A standard estimate of the boundary terms yields

$$\Gamma(z)=\lim_{n\to\infty}|\Lambda_n|^{-1}\log\left[\frac{Z_{\Lambda_n}(U+zX|x(V^c))}{Z_{\Lambda_n}(U|x(V^c))}\right]$$

uniformly in $x(V^c)$. The family of holomorphic functions $z\mapsto|\Lambda_n|^{-1}\log[Z_{\Lambda_n}(U+zX|x(V^c))$ is uniformly bounded on the complex disk $|z|<\varepsilon$ from (2.4). From Montel's theorem this family is normal: the functions in (2.5) can be defined on this disk, the limit Γ is holomorphic, the convergence is uniform on compact subsets of the complex disk $|z| < \varepsilon$, and we can interchange limit and derivatives (see [24], Theorems 14.6 and 10.27). Hence

$$\Gamma'(0) = \lim_n \Gamma'_n(0) = \lim_n \left[\mathbf{E}^{\pi_{\Lambda_n}}(\bar{x}_n)\right](x(V^c))$$

and it follows that all solutions ν of (1.8) have the same asymptotic mean value $\lim_n \mathbf{E}^\nu \bar{x}_n = \Gamma'(0)$. Moreover

$$\begin{aligned} \sigma^2 = \Gamma''(0) &= \lim_n \Gamma''_n(0) \\ &= \lim_n \left[\mathbf{E}^{\pi_{\Lambda_n}}(\bar{x}_n^2) - (\mathbf{E}^{\pi_{\Lambda_n}} \bar{x}_n)^2\right](x(V^c)) \\ &> 0 \end{aligned}$$

since α is not a Dirac mass [8]. Being the Legendre transform of the smooth, strictly convex on $(-\varepsilon, \varepsilon)$ function Γ, I is itself analytic in a neighborhood of $I^{-1}(0)$.

2. We set $\bar{x}_n^o = \bar{x}_n - \mathbf{E}^\nu \bar{x}_n$ to simplify notations. Fix next $a > 0, (\sigma^2 + a)t < 1$. It follows from the uniform convergence of Γ''_n to Γ'' on compact subsets of the disk $|z| < \varepsilon$ that we can find $b > 0$ and $N > 1$ such that $\sup\{|\Gamma''_n(s) - \Gamma''(s)|; |s| < b, n \geq N\} < a$. By Taylor's formula, for arbitrary $\bar{\delta} > 0$, reducing b if necessary it holds that $|\Gamma_n(s) - s\Gamma'_n(0)| \leq s^2(\sigma^2 + a + \bar{\delta})/2$ for $|s| < b$, and the classical Chebycheff bound implies, for $n \geq N$,

$$\begin{aligned} \nu(|\bar{x}_n^o| > u) &\leq 2\exp\left(-|\Lambda_n| \sup\{su - s^2(\sigma^2 + a)/2; s \in (-b, b)\}\right) \\ &= 2\exp\left(-|\Lambda_n| u^2/[2(\sigma^2 + a + \bar{\delta})]\right) \end{aligned} \tag{2.9}$$

for $u < \delta = b(\sigma^2 + a)$. Using again Taylor's formula,

$$\begin{aligned} \mathbf{E}^\nu(\exp\{t|\Lambda_n|^{1/2}(\bar{x}_n^o)\}) &= \exp|\Lambda_n|\{\Gamma_n(t|\Lambda_n|^{-1/2}) - t|\Lambda_n|^{-1/2}\mathbf{E}^\nu \bar{x}_n\} \\ &= \exp|\Lambda_n| \int_0^{t|\Lambda_n|^{-1/2}} (t|\Lambda_n|^{-1/2} - s)\Gamma''_n(s)ds \end{aligned}$$

which converges to $\exp \sigma^2 t^2/2$ as $n \to \infty$. Hence $|\Lambda_n|^{1/2}(\bar{x}_n^o)$ converges in distribution (under the law ν) to a centered Gaussian variable η with variance σ^2, and

$$\exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n^o)^2\}\mathbf{1}_{|\bar{x}_n^o| < \delta} \to \exp t\eta^2/2$$

under ν. The estimate (2.9) shows that the left hand side is uniformly integrable when $\sigma^2 t < 1$, therefore the ν-expectations converge too and the second statement in Proposition 2 is proved. Note also that, again from (2.9),

$$\begin{aligned} \left|\mathbf{E}^\nu\left[\exp\{\frac{z|\Lambda_n|}{2}(\bar{x}_n^o)^2\}\mathbf{1}_{|\bar{x}_n^o| < \delta}\right]\right| &\leq \mathbf{E}^\nu\left[\exp\{\frac{|\Re(z)||\Lambda_n|}{2}(\bar{x}_n^o)^2\}\mathbf{1}_{|\bar{x}_n^o| < \delta}\right] \\ &\leq [1 - (\sigma^2 + a + \bar{\delta})|\Re(z)|]^{-1/2} \end{aligned}$$

provided that $(\sigma^2 + a + \bar{\delta})|\Re(z)| < 1$ and n is large enough. Hence the sequence of holomorphic function $\mathbf{E}^\nu[\exp\{\frac{z|\Lambda_n|}{2}(\bar{x}_n^o)^2\}\mathbf{1}_{|\bar{x}_n^o|<\delta}]$ remains bounded on compact subsets of the above complex domain: by convergence of the derivatives we obtain (2.6).

3. Regardless of the value of t, we saw in the proof of the first part that $\sigma_n \to \sigma$. Turning to the case $\sigma^2 t = 1$, it is easily checked that for $I^{-1}(0)$ to be a local minimum of I_t it is necessary that $\Gamma^{(3)}(0) = 0$. Next we use the identity

$$\exp\{\frac{ta^2}{2|\Lambda_n|}\} = \int_{\mathbb{R}} e^{ay} \left(\frac{|\Lambda_n|}{2\pi t}\right)^{1/2} \exp\{-\frac{|\Lambda_n|y^2}{2t}\} dy$$

to write (with $a = |\Lambda_n|\bar{x}_n^o$)

$$\begin{aligned} Z_n &:= \mathbf{E}^\nu\left(\exp\{\frac{t_n|\Lambda_n|}{2}(\bar{x}_n^o)^2\}\right) \\ &= \left(\frac{|\Lambda_n|}{2\pi t_n}\right)^{1/2} \int_{\mathbb{R}} \exp\{-\frac{|\Lambda_n|y^2}{2t_n}\}\mathbf{E}^\nu\left(\exp\{|\Lambda_n|\bar{x}_n^o y\}\right) dy \\ &= \left(\frac{|\Lambda_n|}{2\pi t_n}\right)^{1/2} \int_{\mathbb{R}} \exp\{|\Lambda_n|(\Gamma_n(y) - y\Gamma_n'(0) - y^2/2t_n)\} dy\,, \qquad (2.10) \end{aligned}$$

by the definition of Γ_n. The last integral will be analyzed by means of the Laplace method. Note that since $\Gamma_n(y)$ increases at most linearly at infinity, it is enough to consider a large enough compact neighborhood of 0 when evaluating the last integral, and on this neighborhood convergence of Γ_n, and further of $\Gamma_n(y) - y\Gamma_n'(0) - y^2/2t_n$, is uniform due to convexity of Γ_n. Since the function $I(y) - \frac{t}{2}(y - I^{-1}(0))^2$ possesses a unique minimizer at $I^{-1}(0)$, it follows by convex duality that $\Gamma(y) - y\Gamma'(0) - y^2/2t$ has 0 as its unique maximum. Therefore, for all $\delta > 0$ one has

$$Z_n = \left(\frac{|\Lambda_n|}{2\pi t_n}\right)^{1/2} \int_{-\delta}^{\delta} \exp\{|\Lambda_n|(\Gamma_n(y) - y\Gamma_n'(0) - y^2/2t_n)\} dy + O(e^{-C(\delta)|\Lambda_n|}) \qquad (2.11)$$

On the other hand, by uniform convergence of $\Gamma_n^{(4)}$ in a neighborhood of the origin, and the fact that $\Gamma^{(4)}(0) < 0$, one has that for δ small enough, there exists a constant $\kappa > 0$ such that, for all $|y| \leq \delta$,

$$-\kappa y^4 \leq \Gamma_n(y) - \Gamma_n'(0)y - \Gamma_n''(0)y^2/2 - \Gamma_n^{(3)}(0)y^3/6 \leq -\kappa^{-1} y^4\,.$$

Substituting in (2.11) and evaluating the integral, recalling that $\Gamma_n^{(3)}(0) = 0$, we observe that

$$C_1|\Lambda_n|^{1/4} \leq Z_n \leq C_2|\Lambda_n|^{1/4}\,,$$

proving (2.7). In order to prove (2.8) we start with a remark. Making explicit the dependence on t_n of the quantity $Z_n = Z_n(t_n)$ defined in (2.10), we have for all fixed $s > 0$

$$Z_n(t_n + s/\sqrt{|\Lambda_n|}) = O(\Lambda_n^{1/4}) \tag{2.12}$$

Indeed, the same estimates as before lead to bound (2.11) from above by

$$C|\Lambda_n|^{1/2} \int_{-\delta}^{\delta} \exp\{-\kappa^{-1}|\Lambda_n|y^4 + C|\Lambda_n|^{1/2}y^2\}dy = O(|\Lambda_n|^{1/4})$$

Here and below C denotes a positive constant, depending on s but not on n, which may change from line to line. Now, from the obvious inequality $u^2 \leq C(1 + \exp su^2)$, one gets with $u = |\Lambda_n|^{1/4}(\bar{x}_n^o)$

$$\mathbf{E}^\nu \left(|\Lambda_n|(\bar{x}_n^o)^2 \ \exp\{\frac{t_n|\Lambda_n|}{2}(\bar{x}_n^o)^2\} \right) \leq C|\Lambda_n|^{1/2} \left(Z_n(t_n + |\Lambda_n|^{-1/2}s) + Z_n(t) \right)$$

which, in view of (2.12), shows (2.8).

4. The proof is very similar. Note that, with $\sigma^2 = \Gamma''(0)$ and $\sigma_n^2 := \Gamma_n''(0)$, we have by translation invariance of ν that

$$\sigma^2 - \sigma_n^2 = \sum_{i \in \mathbb{Z}^d} \text{cov}_\nu(x_0, x_i) - \text{var}_\nu(|\Lambda_n|^{-1/2} \sum_{i \in \Lambda_n} x_i) = |\Lambda_n|^{-1} \sum_{i \in \Lambda_n, j \notin \Lambda_n} \text{cov}_\nu(x_i, x_j)\,.$$

Because U is completely analytic, covariances decay exponentially (see [10], condition II-c), that is $\text{cov}_\nu(x_i, x_j) = O(\exp -C|i-j|)$, and thus $|\Lambda_n|\ |\sigma^2 - \sigma_n^2| \leq C|\partial\Lambda_n| = C|\Lambda_n|^{1/2}$ in two dimensions. A similar computation, using again [10], condition II-c, shows also that $|\Lambda_n|\ |\Gamma_n^{(3)}(0) - \Gamma^{(3)}(0)| \leq C|\Lambda_n|^{1/2}$. This time the substitution in the integral term of (2.11) leads to upper and lower bounds

$$\left(\frac{|\Lambda_n|}{2\pi t}\right)^{1/2} \int_{-\delta}^{\delta} \exp\{-\kappa^{\mp 1}|\Lambda_n|y^4 \pm C|\Lambda_n|^{1/2}(y^2 + |y|^3)\}dy = O(|\Lambda_n|^{1/4})$$

proving (2.7) when t replaces t_n in dimension $d = 2$. Combined with (2.12) the same arguments clearly conclude to the analoguous version of (2.8) in this case. □

Remark 2 *We consider briefly here the case where the infinite volume measure ν is replaced in the previous Proposition with a sequence of finite volume ν_n satisfying the specification π (with appropriate boundary conditions). An inspection of the proof reveals that all statements remain valid with ν_n instead of ν. We emphasize that in this case the centering in (2.6, 2.7, 2.8) becomes $\mathbf{E}^{\nu_n}\bar{x}_n$, and that in point 3, Γ_n is defined by (2.5) with ν_n instead of ν, so that $t_n = (var_{\nu_n}(|\Lambda_n|^{-1/2}\sum_{i\in\Lambda_n} x_i))^{-1}$.*

We next turn our attention to the case where the perturbation shifts the mean field. In anticipation of the applications to (finite volume) Gibbs fields, we let ν_n denote a sequence of measures on Σ^{Λ_n} corresponding to a translation invariant potential U, and we denote by ν its infinite volume limit. We will assume below that $\mathbf{E}^{\nu_n}\bar{x}_n = 0$ for all n. Obviously, in this case, $I^{-1}(0) = 0$.

A crucial role will be played in our investigation by the set of minimizers

$$\mathcal{M}_t = \{\ y : I_t(y) = \min I_t\} \tag{2.13}$$

of the functions $I_t(y) : y \mapsto -\frac{t}{2}(y - I^{-1}(0))^2 + I(y) = -ty^2/2 + I(y)$. We let $\bar{I}_t := \inf_y I_t(y) = I_t(y^*)$ for all $y^* \in \mathcal{M}_t$. We say that $y^* \in \mathcal{M}_t$ is *nondegenerate* if I_t is twice continuously differentiable at y^* and $I_t''(y^*) = -t + I''(y^*) > 0$. Note that in the context of Proposition 2, when U is completely analytic for the mean, the condition $\sigma^2 t > 1$ corresponds to 0 being a local maximizer (and not minimizer) of the function $I_t(\cdot)$.

We need to introduce the analogue of I_t and of $\mathcal{M}_t$ in the case ν_n replaces ν. That is, define

$$\Gamma_n(t) = |\Lambda_n|^{-1} \log \mathbf{E}^{\nu_n} e^{t|\Lambda_n|\bar{x}_n}\,, \quad I^n(y) = \sup_{t\in\mathbb{R}}\{ty - \Gamma_n(t)\}\,, \quad I_t^n(y) = -\frac{t}{2}y^2 + I^n(y)$$

and let $\mathcal{M}_t^n$ denote the set of minimizers of $I_t^n(\cdot)$, with $\bar{I}_t^n := \inf_y I_t^n(y)$.

We denote by ν_n^h the measure on Σ^{Λ_n} corresponding to the potential U^h, i.e

$$\frac{d\nu_n^h}{d\nu_n} := \frac{\exp\{|\Lambda_n|h\bar{x}_n\}}{Z_n^h}\,, \quad Z_n^h := \mathbf{E}^{\nu_n}\exp\{|\Lambda_n|h\bar{x}_n\} = \exp|\Lambda_n|\Gamma_n(h)\,. \tag{2.14}$$

Let $h_n(y)$ be defined such that $\mathbf{E}^{\nu_n^{h_n(y)}}(\bar{x}_n) = y$.
It is convenient to make use of the following assumption, which could be called "complete analyticity for the mean". See Remark 4 following Proposition 5 for comments on this assumption.

Assumption (CAM): U and $\{\nu_n\}$ are symmetric with respect to the transformation $x \mapsto -x$ in Σ^{Λ_n}. $\mathcal{M}_t$ consists of two elements $\pm y$. Let h be the unique point where the subdifferential of the convex function Γ contains y, and assume that $U^h = U + hX$ is completely analytic for the mean.

Proposition 3 *Assume* **(CAM)**.

1. *Γ is analytic on a complex disc of radius ϵ around h. I_t is analytic on an ϵ neighborhood of $\mathcal{M}_t$.*

2. *Assume that* $\Gamma''(h)t < 1$, *i.e. that* $\pm y$ *is a nondegenerate minimum of* $I_t(\cdot)$. *Then for* n *large enough there exist exactly two minimizers* $\pm y_n$ *of* $I_t^n(\cdot)$. *Further, there exists a constant* δ *independent of* n *such that, as* $n \to \infty$,

$$\lim_n \exp\{|\Lambda_n|\bar{I}_t^n\}\mathbf{E}^{\nu_n} \exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\} = 2(1-\Gamma''(h)t)^{-1/2} \tag{2.15}$$

and

$$\lim_n e^{|\Lambda_n|\bar{I}_t^n}\mathbf{E}^{\nu_n}\left(|\Lambda_n|(\bar{x}_n - y_n)^2\, e^{\frac{t|\Lambda_n|}{2}\bar{x}_n^2}\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right) = \frac{\Gamma''(h)}{(1-\Gamma''(h)t)^{3/2}}\,. \tag{2.16}$$

Finally,

$$\limsup_{n\to\infty}\frac{1}{|\Lambda_n|}\log\left(\exp\{|\Lambda_n|\bar{I}_t^n\}\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\}\mathbf{1}_{\bar{x}_n\notin(\mathcal{M}_t)^\delta}\right)\right) < 0\,. \tag{2.17}$$

Proof: 1. The proof of part 1 is a rerun of the argument used in Proposition 2, part 1. Note in addition that $\theta_n \to y$ implies that $h_n(\theta_n) \to h$, since $\Gamma_n'(h_n(\theta_n)) = \theta_n$ and $\Gamma_n' \to \Gamma'$ uniformly in a neighborhood of h with $\Gamma'(h) = y$.
2. Recall that $\Gamma_n \to \Gamma$ on the whole real line, so $I^n \to I$ by convex duality. The nondegeneracy condition implies that the second derivative of I_t is strictly positive in a neighborhood of $\pm y$. Due to Montel's theorem, c.f. the proof of Theorem 1, this property is inherited by $I_t^n(\cdot)$. Therefore, the latter function possesses also only two minima, which due to symmetry are $\pm y_n$. The corresponding $h_n := h_n(y_n)$ satisfy $h_n \to h$ and $h_n(-y_n) = -h_n(y_n)$.

Recall next that our perturbation $t(\cdot)^2/2$ is a bounded and continuous function. By Varadhan's lemma (see, e.g., for this application [5, Exercise 4.3.11]) and the existence of large deviations for $\bar{x}_n$ under ν_n irrespective of the boundary conditions, it holds that

$$\limsup_{n\to\infty}\frac{1}{|\Lambda_n|}\log\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\}\mathbf{1}_{\bar{x}_n\notin(\mathcal{M}_t)^\delta}\right) \le -\inf_{y\notin(\mathcal{M}_t)^\delta} I_t(y) < 0\,,$$

where the last inequality is due to the lower semicontinuity of $I_t(\cdot)$. This proves (2.17).
Note next that by definition,

$$(I^n)'(y_n) = ty_n = h_n\,,\ \ \Gamma_n(h_n) = h_n y_n - I^n(y_n)\,. \tag{2.18}$$

Hence,

$$\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\}\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right)$$

$$= \exp\{-\frac{t|\Lambda_n|}{2}y_n^2\}\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n - y_n)^2 + t|\Lambda_n|y_n\bar{x}_n\}\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right)$$
$$= \exp\{-\frac{t|\Lambda_n|}{2}y_n^2 + |\Lambda_n|\Gamma_n(ty_n)\}\mathbf{E}^{\nu_n^{h_n}}\left(\exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n - y_n)^2\}\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right)$$
$$= \exp\{-|\Lambda_n|I_t^n(y_n)\}\mathbf{E}^{\nu_n^{h_n}}\left(\exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n - y_n)^2\}\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right) \tag{2.19}$$

by (2.18). Repeating now the argument in the proof of part 2 of Proposition 2, using this time the potential U^{h_n} and the fact that y_n is the unique minimum and is nondegenerate, one concludes that $(\bar{x}_n - y_n)\sqrt{|\Lambda_n|}$, localized to $\{|\bar{x}_n - y_n| < \delta\}$, converges under $\nu_n^{h_n}$ to the Gaussian distribution $\mathcal{N}(0, \Gamma''(h))$. Uniform integrability of $\exp\{\frac{t|\Lambda_n|}{2}(\bar{x}_n - y_n)^2\}$ on $\{|\bar{x}_n - y_n| < \delta\}$ for small δ follows from uniform convergence of I^n to I in a neighborhood of y and nondegeneracy. Hence the last expectation in (2.19) tends to $(1 - \Gamma''(h)t)^{-1/2}$. In addition to (2.17) and symmetry, this implies (2.15). The proof of (2.16) is analogous.

□

3 Mean-field perturbations of Gibbs random fields

Let π denote a specification, and let ν_n denote a sequence of measures on Λ_n satisfying the specification π (with appropriate boundary conditions). When all measures ν_n are restrictions to Λ_n of the same Gibbs measure ν on $\Sigma^{\mathbb{Z}^d}$, we use ν to denote also ν_{Λ_n}. For simplicity we assume uniqueness of the solution ν on $\Sigma^{\mathbb{Z}^d}$ of (1.8), which is the case when π is weak-mixing, and we also assume that $\mathbf{E}^{\nu}x_0 = 0$. For $t > 0$, and for a sequence of measures ν_n, consider the sequence of probability measures $\mu_n = \mu_{n,t}$ on $\Sigma^{\mathbb{Z}^d}$,

$$d\mu_n = Z_n^{-1}\exp\{|\Lambda_n|t\bar{x}_n^2/2\}\, d\nu_n\,, \tag{3.1}$$
$$Z_n = \mathbf{E}^{\nu_n}\exp\{|\Lambda_n|t\bar{x}_n^2/2\}\,. \tag{3.2}$$

In this section we deduce fine asymptotics for μ_n from our previous estimates. We deal first with the case $\nu_n = \nu_{\Lambda_n}$ and $\mathcal{M}_t = \{0\}$ (recall (2.13)).

Proposition 4 *Let π satisfy (2.4), assume the associated Gibbs measure ν satisfies $\mathcal{M}_t = \{0\}$, and let $\sigma^2 = \Gamma''(0) = I''(0)^{-2}$. Define $\mu_n = \mu_{n,t}$ by (3.1) with $\nu_n = \nu_{\Lambda_n}$.*

1. *If in addition*

$$y = 0 \text{ is non-degenerate, i.e. } t < \sigma^{-2}\,, \tag{3.3}$$

then

$$\lim_{n\to\infty} H_{\Lambda_n}(\mu_n|\nu) = (1/2)\{\sigma^2 t/(1-\sigma^2 t) + \log(1-\sigma^2 t)\}$$

2. *When $\sigma^2 t = 1$, let $\sigma_n^2 = \Gamma_n''(0)$ and $t_n = \sigma_n^{-2}$. Assume further that $\Gamma_n^{(3)}(0) = 0, \Gamma^{(4)}(0) < 0$. Then*

$$\limsup_n |\Lambda_n|^{-1/2} H_{\Lambda_n}(\mu_{n,t_n}|\nu) < \infty$$

3. *In dimension $d = 2$, if $\sigma^2 t = 1, \Gamma^{(4)}(0) < 0$ and U is completely analytic, we have also*

$$\limsup_n |\Lambda_n|^{-1/2} H_{\Lambda_n}(\mu_{n,t}|\nu) < \infty$$

Proof: By definition we have

$$H_{\Lambda_n}(\mu_n|\nu) = \frac{t|\Lambda_n|}{2}\mathbf{E}^{\mu_n}\bar{x}_n^2 - \log Z_n\,, \tag{3.4}$$

where we recall that $Z_n = \mathbf{E}^\nu \exp\{|\Lambda_n| t\bar{x}_n^2/2\}$.
1) Note that $\mathbf{E}^\nu \bar{x}_n = \mathbf{E}^\nu x_0 = 0$ follows from the uniqueness of ν. We have from Proposition 2 that

$$Z_n^\delta = \mathbf{E}^\nu \left(\exp\{|\Lambda_n| t\bar{x}_n^2/2\}\mathbf{1}_{|\bar{x}_n|<\delta}\right) \sim (1-\sigma^2 t)^{-1/2}\,,$$

though from large deviations

$$\limsup_{n\to\infty} |\Lambda_n|^{-1}\log(Z_n - Z_n^\delta) \le -\max\{ty^2/2 - I(y)\ ;\ |y| \ge \delta\} < 0$$

since $\mathcal{M}_t = \{0\}$. Hence $Z_n \sim (1-\sigma^2 t)^{-1/2}$. We decompose in a similar manner

$$Z_n \mathbf{E}^{\mu_n}\left(|\Lambda_n|\bar{x}_n^2\right) = Y_n^\delta + (Y_n - Y_n^\delta)$$

$$Y_n^\delta = \mathbf{E}^\nu \left(|\Lambda_n|\bar{x}_n^2 \exp\{|\Lambda_n| t\bar{x}_n^2/2\}\mathbf{1}_{|\bar{x}_n|<\delta}\right) \quad , \quad Y_n = Y_n^\infty$$

Large deviations then imply that $Y_n - Y_n^\delta$ is exponentially small, though $Y_n^\delta \sim \sigma^2(1-\sigma^2 t)^{-3/2}$ from Proposition 2. Inserting the equivalents of Y_n, Z_n into (3.4) we obtain the desired result.
2 and 3) The proof of the other statements uses the two last results of Proposition 2 in a similar manner, so details are left to the reader. See [1] for a similar argument. □

We briefly consider the case when ν is replaced in (3.1) with a finite volume Gibbs measure ν_n on Λ_n, with $\mathbf{E}^{\nu_n}\bar{x}_n = 0$ to get the proper centering; we have in mind the case of free or periodic boundary conditions. The estimates mentioned in Remark 2 lead in a similar fashion to:

Remark 3 *The limits in the previous Proposition still hold under the set of assumptions therein, when ν is replaced in (3.1) with a sequence of measures ν_n on Λ_n satisfying the specification π (with appropriate boundary conditions), provided that $\mathbf{E}^{\nu_n}\bar{x}_n = 0$ for all n and that all informations $(H_{\Lambda_n}(\mu_n|\nu_n),..)$ are computed with respect to ν_n. We emphasize that in point 2, t_n takes on the value $t_n = (var_{\nu_n}(|\Lambda_n|^{-1/2}\sum_{i\in\Lambda_n} x_i))^{-1}$.*

As in [1], one may encounter mixtures when the perturbation is not small anymore, leading to multiple minima.

Proposition 5 *Fix $t > 0$ and assume* **CAM**. *With the notations of Proposition 3, in particular $h_n = h_n(y_n)$, let*

$$\bar{\nu}_t^n := \frac{\nu_n^{h_n} + \nu_n^{-h_n}}{2}.$$

Then,

$$\limsup_{n\to\infty} H_{\Lambda_n}(\mu_n|\bar{\nu}_t^n) < \infty. \tag{3.5}$$

Further (c.f. the notations of Theorem 1), with $m/n \to 0$, $m \to \infty, n \to \infty$,

$$|A(m,n)|^{-1} \sum_{j\in A(m,n)} \|\mu_n - \bar{\nu}_t^n\|^2_{j+\Lambda_m} \to_{n\to\infty} 0. \tag{3.6}$$

provided that the specification associated with the potential $U + h'X$ is weak-mixing in a neighborhood of h with constants γ, C independent of h'.

Remark 4 *Assuming* **CAM** *is not enough to get (3.6), so we assume the above locally uniform weak-mixing property around h. When $U^h = U + hX$ is completely analytic or when strong mixing (condition III-c in [10]) holds for $U^{h'}$ locally uniformly around h, this property is satisfied.*

Proof: Recall the definition (2.14) and note that $Z_n^{h_n} = Z_n^{-h_n}$ by symmetry. By definition we have

$$\begin{aligned}
H_{\Lambda_n}(\mu_n|\bar{\nu}_t^n) &= \frac{t|\Lambda_n|}{2}\mathbf{E}^{\mu_n}\bar{x}_n^2 - \log Z_n - \mathbf{E}^{\mu_n}\left(\log\left(\frac{1}{2}\left(\frac{d\nu_n^{h_n}}{d\nu_n} + \frac{d\nu_n^{-h_n}}{d\nu_n}\right)\right)\right) \\
&= \frac{t|\Lambda_n|}{2Z_n}\mathbf{E}^{\nu_n}\left(\bar{x}_n^2 \exp\{\frac{t|\Lambda|_n}{2}\bar{x}_n^2\}\right) - \log\left(\frac{Z_n}{Z_n^{h_n}}\right) \\
&\quad -\frac{1}{Z_n}\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\}\log\cosh(|\Lambda_n|h_n\bar{x}_n)\right) \\
&= \frac{t|\Lambda_n|}{Z_n}\mathbf{E}^{\nu_n}\left(\bar{x}_n^2 \exp\{\frac{t|\Lambda|_n}{2}\bar{x}_n^2\}\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right) - \log\left(\frac{Z_n}{Z_n^{h_n}}\right) + \log 2 \\
&\quad -\frac{2}{Z_n}\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\}|\Lambda_n|h_n\bar{x}_n\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right) + O(\exp -C|\Lambda_n|),
\end{aligned} \tag{3.7}$$

where the last equality is due to symmetry and the exponential bound (2.17). By definition of ν_n^h and $h_n = ty_n$, c.f. (2.14) and (2.18), we get

$$\begin{aligned} \log\left(\frac{Z_n}{Z_n^{h_n}}\right) &= \log\left(\frac{\mathbf{E}^\nu(\exp\{|\Lambda_n|t\bar{x}_n^2/2\})}{Z_n^{h_n}}\right) \\ &= \log \mathbf{E}^{\nu_n^{h_n}}\left(\exp\{|\Lambda_n|(-h_n\bar{x}_n + t\bar{x}_n^2/2)\}\right) \\ &= \log \mathbf{E}^{\nu_n^{h_n}}\left(\exp\{\frac{t|\Lambda_n|}{2}[(\bar{x}_n - y_n)^2 - y_n^2]\}\right) \\ &\geq -\frac{t|\Lambda_n|}{2}y_n^2 \end{aligned}$$

Substituting in (3.7) and using again (2.18) and the fact that $Z_n^{-1}\mathbf{E}^{\nu_n}(\mathbf{1}_{|\bar{x}_n - y_n|<\delta}) = 1/2 + O(\exp(-C|\Lambda_n|))$, we have

$$\begin{aligned} H_{\Lambda_n}(\mu_n|\bar{\nu}_t^n) &\leq C + \frac{|\Lambda_n|}{Z_n}\mathbf{E}^{\nu_n}\left(\exp\{\frac{t|\Lambda_n|}{2}\bar{x}_n^2\}t(\bar{x}_n - y_n)^2\mathbf{1}_{|\bar{x}_n - y_n|<\delta}\right) \\ &\leq C \end{aligned} \tag{3.8}$$

where the last inequality is due to (2.15) and (2.16). This completes the proof of (3.5).

The proof of the second part hinges upon a localization procedure. Define, on Σ^{Λ_n}, the measures $\mu_n^{\pm\delta} := \mu_n(\,\cdot\,|\,|\bar{x}_n - \pm y_n(t)| < \delta)$. By using the estimate (3.5) and the large deviations, c.f. [1, Theorem 2] for a similar argument, one has that

$$H_{\Lambda_n}(\mu_n^{\pm\delta}|\nu_n^{\pm h_n}) = O(1) \tag{3.9}$$

Indeed, with the notation $A_{n,\pm\delta} = \{|\bar{x}_n - \pm y_n(t)| < \delta\}$,

$$\begin{aligned} H_{\Lambda_n}(\mu_n^{\pm\delta}|\nu_n^{\pm h_n}) &= -\log\mu_n(A_{n,\pm\delta}) + \int\log\left(\frac{d\mu_n}{d\nu_n^{\pm h_n}}\right)d\mu_n^{\pm\delta} \\ &= -\log\mu_n(A_{n,\pm\delta}) + \int\log\left(\frac{d\bar{\nu}_t^n}{d\nu_n^{\pm h_n}}\right)d\mu_n^{\pm\delta} + \int\log\left(\frac{d\mu_n}{d\bar{\nu}_t^n}\right)d\mu_n^{\pm\delta} \\ &= -\log\mu_n(A_{n,\pm\delta}) + \int\log\left(\frac{1+e^{\mp 2h\bar{x}_n|\Lambda_n|}}{2}\right)\mathbf{1}_{A_{n,\pm\delta}}d\mu_n^{\pm\delta} \\ &\quad +\mu_n(A_{n,\pm\delta})^{-1}\int\log\left(\frac{d\mu_n}{d\bar{\nu}_t^n}\right)\mathbf{1}_{A_{n,\pm\delta}}d\mu_n \\ &= (1+o(1))H_{\Lambda_n}(\mu_n|\bar{\nu}_t^n) + 0(1)\,. \end{aligned}$$

By assumption the measures $\nu_n^{h_n}$ are weakly mixing with constants γ, C independent of n, so one concludes from (3.9) and Theorem 1 that in fact, with n, m as in the statement of the proposition,

$$|A(m,n)|^{-1} \sum_{j \in A(m,n)} \| \mu_n^{\pm,\delta} - \nu_n^{\pm h_n} \|_{j+\Lambda_m}^2 \to_{n\to\infty} 0 \,. \tag{3.10}$$

The conclusion (3.6) follows now by observing that

$$\begin{aligned} \|\mu_n - \bar\nu_t^n\|_{j+\Lambda_m} &\le \mu_n\{\bar x_n \notin A_{n,\pm\delta}\} + \|\mu_n\{\bar x_n \in A_{n,\pm\delta}\} \frac{\mu_n^{+\delta} + \mu_n^{-\delta}}{2} - \frac{\nu_n^{h_n} + \nu_n^{-h_n}}{2} \|_{j+\Lambda_m} \\ &= \mu_n\{\bar x_n \notin A_{n,\pm\delta}\} + \|\mu_n\{\bar x_n \in A_{n,\pm\delta}\} \mu_n^{+\delta} - \nu_n^{h_n} \|_{j+\Lambda_m} \\ &\le 2\mu_n\{\bar x_n \notin A_{n,\pm\delta}\} + \mu_n\{\bar x_n \in A_{n,\pm\delta}\} \|\mu_n^{+\delta} - \nu_n^{h_n}\|_{j+\Lambda_m} \end{aligned}$$

□

Remark 5 *The symmetry played an essential role in our proof. Indeed, without it, it would have been impossible to relate the value of Z_n and $Z_n^{h_n}$ in a straightforward way.*

Example 1. We consider the ferromagnetic Ising model without external field: let Σ be the two-point set $\{-1, +1\}$, α be the fair Bernoulli measure, let

$$U_A(x) = \beta x_i x_j \qquad \text{if } A = \{i,j\} \text{ with } i,j \text{ nearest neighbors}$$
$$U_A = 0 \quad \text{otherwise} \tag{3.11}$$

and denote by β_c the critical value. When $\beta \le \beta_c$ the solution ν of (1.8, 2.2) is unique, invariant under translations and under the global flip $x \mapsto -x$. We will consider also the potential $U^h = U + hX$ with external field h. In dimension $d = 2$ complete analyticity for U^h has been proved in the full uniqueness region $\beta < \beta_c$ or $h \neq 0$ [25]; in higher dimension, weak-mixing for U^h is known in the strips $\beta < \beta_c$ [15] and $\beta > \beta_1, h \neq 0$ for some $\beta_1 > \beta_c$, though complete analyticity is known for $\beta < \beta_2$ ($\beta_2 < \beta_c$) and strong mixing (condition III-c in [10]) is known to hold (locally) uniformly for large βh [21].
Consider μ_n defined (without external field) by (3.1), $t > 0$ and $\nu_n = \nu_{\Lambda_n}$ for the moment, in the following cases.

case a) $\beta < \beta_c$, $t < \sigma^{-2} = (\sum_i \mathbf{E}^\nu x_0 x_i)^{-1}$ and U completely analytic. Then,

$$I(y) \ge y^2/(2\sigma^2) \,. \tag{3.12}$$

This can be seen from Theorem 3 in Newman [23], which implies by series expansion the following sub-Gaussian property

$$\mathbf{E}^\nu \exp t|\Lambda_n|^{-1/2}\bar x_n \le \mathbf{E} \exp t\eta \;\;,\;\; \eta \sim \mathcal{N}(0,\sigma^2)\,,$$

which leads to $\Gamma(t) \leq \sigma^2 t^2/2$, and then to (3.12). When t is strictly smaller than the inverse susceptibility σ^{-2} the assumption (3.3) is fulfilled, and therefore Proposition 4, Theorem 1 and Corollaries 1, 2 are in force: the perturbed measure μ_n is very close to the original Ising measure ν on cubes whose size vanishes with respect to the volume.

case b) $\beta < \beta_c$, $t\sigma^2 = 1$ and U completely analytic. We define $t_n = \sigma_n^{-2}$ as in Proposition 4. Then part 2 of Proposition 4 applies: indeed, by symmetry we have that $\Gamma_n^{(3)}(0) = 0$ while we have that

$$\Gamma^{(4)}(0) = \lim_{n\to\infty} \Gamma_n^{(4)}(0) =$$

$$\sum_{i,j,k\in\mathbb{Z}^d} (\mathbf{E}^\nu(x_0x_ix_jx_k) - [\mathbf{E}^\nu(x_0x_i)\mathbf{E}^\nu(x_jx_k)+\mathbf{E}^\nu(x_0x_j)\mathbf{E}^\nu(x_ix_k)+\mathbf{E}^\nu(x_0x_k)\mathbf{E}^\nu(x_ix_j)]) \,.$$

By the Gaussian domination [23], all summands in the square brackets above are non-positive. Moreover, for $i = j = k = 0$ the bracket takes the value -2. Hence, $\Gamma^{(4)}(0) \leq -2 < 0$. Since Γ' is concave on $(0,\infty)$ from the GHS inequality [12] (see the argument below in case c)), it follows that $\Gamma''(u) \leq \Gamma''(0^+) = \sigma^2$, and hence that $\Gamma(u) < \sigma^2 u^2/2, (u \neq 0)$. Therefore, by duality, $I_t(y)$ achieves its unique minimum at 0. It follows from part 2 of Proposition 4 that

$$H_{\Lambda_n}(\mu_{n,t_n}|\nu) = O(|\Lambda_n|^{1/2}) \,.$$

Theorem 1 then applies and yields that the measure μ_{n,t_n} is close to ν in the variation norm on boxes translates of Λ_m with $m = o(n^{1/2})$. On the other hand in two dimensions it follows from part 3 of Proposition 4 that the same results hold for $\mu_{n,t}$.

case c) $\beta < \beta_c$, $t\sigma^2 > 1$ and U completely analytic [respectively $\beta \geq \beta_c$, $t > 0$, but U^h completely analytic, where h is the positive maximizer of $h^2/2 - \Gamma(h)$, see below]. We need here the measure ν_n appearing on the right hand side of (3.1) to be symmetric: it will be the finite volume Gibbs measure with either free or periodic boundary conditions, or possibly the infinite volume measure in the case $\beta \leq \beta_c$. Obviously, in this case 0 is a local maximizer of $I_t(\cdot)$. Due to the symmetry of $I_t(y)$, it is enough to discuss $y \geq 0$ [respectively $y \geq y^+ = \mathbf{E}^\nu x_0$], and to show $I_t(y)$ possesses a unique minimum on this interval. This amounts to show that, for the conjugate functions, $h^2/2 - \Gamma(h)$ has a unique minimum on $[0,\infty)$. The GHS inequality [12] states that

$$\begin{aligned}\Gamma_n'''(h) \;=\; & |\Lambda_n|^{-1} \sum_{i,j,k\in\Lambda_n} \mathbf{E}^{\nu_n^h}(x_kx_ix_j) \\ & -|\Lambda_n|^{-1} \sum_{i,j,k\in\Lambda_n} [\mathbf{E}^{\nu_n^h}(x_kx_i)\mathbf{E}^{\nu_n^h}(x_j)+\mathbf{E}^{\nu_n^h}(x_kx_j)\mathbf{E}^{\nu_n^h}(x_i)+\mathbf{E}^{\nu_n^h}(x_ix_j)\mathbf{E}^{\nu_n^h}(x_k)]\end{aligned}$$

$$+2|\Lambda_n|^{-1}\sum_{i,j,k\in\Lambda_n}\mathbf{E}^{\nu_n^h}(x_k)\mathbf{E}^{\nu_n^h}(x_i)\mathbf{E}^{\nu_n^h}(x_j)\ \le\ 0$$

for $h \ge 0$. But for the Ising model it is well known (c.f., for example, [12]) that Γ is a C^1 function on $h > 0$, and by convexity $\Gamma' = \lim \Gamma'_n$ on this interval. Hence Γ' is itself concave on $h > 0$ and uniqueness of the minimizer follows. Therefore, since U^h is completely analytic, Propositions 3 and 5, and Theorem 1, apply and show that the measure μ_n is close to the mixture $[\nu_n^{h_n} + \nu_n^{-h_n}]/2$ (where $h_n = h_n(m_n)$) in the variation norm on boxes translates of Λ_m with $m = o(n)$.

Turning now more systematically to the case of a *sequence* of finite volume Gibbs measures ν_n in (3.1), we can use Remark 3 in the cases a) and b) to get similar conclusions as above. In dimension $d > 3$ we would require symmetry of the ν_n's to ensure $\Gamma_n^{(3)}(0) = 0$. The case c) has been already discussed above. In particular, in the two-dimensional case with periodic boundary conditions we obtain all the results claimed in the Introduction.

We conclude this section with some additional comments. We have recently learned from F. Martinelli [18] that under weak mixing conditions, with arbitrary boundary conditions on $\partial\Lambda_n$ one has that with any $\Delta_n \subset \Lambda_n$ with $d(\Delta_n, \partial\Lambda_n)/\log n \to \infty$,

$$||\nu_n(\cdot|\bar{x}_n = y, x(\Lambda_n^c)) - \nu_n^{h_n(y)}(\cdot|x(\Lambda_n^c))||_{\Delta_n} = O(|\Delta_n|/|\Lambda_n|)\,. \tag{3.13}$$

Integrating over the conditioning and taking into account the fluctuations of $\bar{x}_n$, one checks that this implies, in the context of part 1 of Proposition 4, that for any f with $|f| \le 1$, and with $\bar{\mu}_n$, $\bar{\nu}_n$ denoting the law of $\bar{x}_n$ under μ_n, ν_n,

$$\begin{aligned}
&\int_{\Delta_n} f(x)[\mu_n(dx) - \nu_n(dx)] = \\
&\int\int_{\Delta_n} f(x)[\nu_n(dx|\bar{x}_n = y)\bar{\mu}_n(dy) - \nu_n(dx|\bar{x}_n = y)\bar{\nu}_n(dy)] \\
&= \int\int_{\Delta_n} f(x)[\nu_n(dx|\bar{x}_n = y) - \nu_n^{h_n(y)}(dx)][\bar{\mu}_n(dy) - \bar{\nu}_n(dy)] \\
&\quad + \int\int_{\Delta_n} f(x)\nu_n^{h_n(y)}(dx)[\bar{\mu}_n(dy) - \bar{\nu}_n(dy)] \\
&= O(\frac{|\Delta_n|}{|\Lambda_n|}) + \int\int_{\Delta_n} f(x)[\nu_n^{h_n(y)}(dx) - \nu_n(dx)][\bar{\mu}_n(dy) - \bar{\nu}_n(dy)] \\
&\le O(\frac{|\Delta_n|}{|\Lambda_n|}) + \int ||\nu_n^{h(y)} - \nu_n||_{\Delta_n}[\bar{\mu}_n(dy) + \bar{\nu}_n(dy)]\,,
\end{aligned}$$

where all the estimates above are uniform in f. (In the previous display, we continue to denote, when no confusion occurs, the restriction of μ_n, ν_n on Δ_n by the same notation). A direct computation reveals that $H(\nu_n|\nu_n^{h(y)}) = O(|\Lambda_n|h^2(y))$, while under both $\bar{\mu}_n$ and $\bar{\nu}_n$, $h(\bar{x}_n)|\Lambda_n|^{1/2}$ converge to non-degenerate Gaussian random variables. Hence, using Corollary 2, one concludes that

$$||\mu_n - \nu||_{\Delta_n} = O(\sqrt{|\Delta_n|/|\Lambda_n|}) + O(|\Delta_n|^{-1/d}), \qquad (3.14)$$

namely the same order as we get above, for large $|\Delta_n|$. Note that sharp conditioning results with a faster rate of convergence than the action of mean field perturbations, see the different rates in the right hand sides of (3.13) and (3.14).

We also note some recent results of K. Marton [22] concerning concentration inequalities for (strongly) mixing Markov fields. We do not see however how to derive our results from hers.

References

[1] Ben Arous, G.; Zeitouni, O.: *Increasing propagation of chaos for mean field models.* To appear, Ann. Inst. H. Poincaré (1998)

[2] Bodineau, T.; Presutti, E.: *Phase diagram of Ising systems with additional long range forces.* Comm. Math. Phys. **189**, 287 (1997)

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

[4] Dembo, A.; Zeitouni, O.: *Refinements of the Gibbs conditioning principle.* Prob. Th. Rel. Fields. **104**, 1 (1996)

[5] Dembo, A.; Zeitouni, O.: Large deviations techniques and applications, Second edition, New York: Springer 1998.

[6] Dembo, A.; Zeitouni, O.: *Refinements of the Gibbs conditioning principle* , unpublished manuscript. Expanded version of [4].
http://www-ee.technion.ac.il/~zeitouni/ps/sanovreffull.ps

[7] Deuschel, J. D.; Stroock, D. W.: Large Deviations. Boston: Academic Press 1989

[8] Dobrushin, R. L.; Nahapetian, B. S.: *Strong convexity of the pressure for lattice systems of classical statistical physics (in Russian).* Teoret. Mat. Fiz. **20**, 223–234, 1974.

[9] Dobrushin, R., Shlosman, S.: *Constructive criterion for the uniqueness of Gibbs fields.* In "Statistical physics and dynamical systems"(Fritz, Jaffe, Szász eds), 347, Birkhauser 347 (1985)

[10] Dobrushin, R.L., Shlosman, S.: *Completely Analytical Gibbs Fields.* In "Statistical physics and dynamical systems"(Fritz, Jaffe, Szász eds), 371, Birkhauser 347 (1985)

[11] Dupuis, P.; Ellis, R.: A weak convergence approach to the theory of large deviations. New York: Wiley (1997)

[12] Ellis, R.S.: Entropy, large deviations, and statistical mechanics. Berlin: Springer (1985)

[13] Föllmer H.: *On entropy and information gain in random fields.* Z. Wahrsch. verw. Gebiete **26**, 207-217 (1973)

[14] Georgii H. O.: Gibbs Measures and Phase Transitions. Berlin: De Gruyter 1988

[15] Higuchi, Y.: *Coexistence of infinite (*)-clusters II: Ising percolation in two dimensions.* Prob. Th. Rel. Fields **97**, 1 (1993)

[16] Lebowitz, J. L.; Penrose, O.: *Rigorous treatment of the Van der Waals-Maxwell theory of liquid vapor transition.* J. Math. Phys. **7**, 98-113 (1966)

[17] Lewis, J. T.; Pfister, C. E.; Sullivan, W. G.: *Entropy, concentration of probability, and conditional limit theorems*, Markov Proc. Rel. Fields **1**, 319–386 (1995).

[18] Martinelli, F. and Cancrini, N.: *Finite volume equivalence of ensembles*, in preparation.

[19] Martinelli, F.; Olivieri, E.; Schonmann, R.: *For Gibbs state of 2D lattice spin systems weak mixing implies strong mixing.* Comm. Math. Phys. **165**, 33 (1994)

[20] Martinelli, F.: *Lectures on Glauber dynamics for discrete spin systems.* Ec. de Probab. St Flour 1997. To appear Lect. N. Math., Springer

[21] Martinelli, F.; Olivieri, E.: *Approach to equilibrium of Glauber dynamics in the one-phase region I: The attractive case.* Comm. Math. Phys. **161**, 447 (1994)

[22] Marton, K.: *On a measure concentration inequality of Talagrand for dependent random variables*, preprint (1998).

[23] Newman, C. : *Gaussian correlation inequalities for ferromagnets.* Z. Wahrsch. verw. Gebiete **33**, 75–93 (1975)

[24] Rudin, W.: Real and complex analysis. Mc Graw-Hill, New York, 2nd ed. 1974

[25] Schonmann, R.; Shlosman, S.: *Complete analyticity for 2D Ising model completed.* Comm. Math. Phys. **170**, 453 (1995)

[26] Schroeder, C.: *I-projection and conditional limit theorems for discrete parameter Markov processes.* Annals Probab. **21**, 721–758, 1993.