Quasi–factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields

Filippo Cesi

Dipartimento di Fisica, Università di Roma "La Sapienza", P.le A. Moro 2, 00185 Roma, Italy
and INFM Unità di Roma "La Sapienza".
cesi@zephyrus.roma1.infn.it

Abstract. We show that the entropy functional exhibits a quasi–factorization property with respect to a pair of weakly dependent σ–algebras. As an application we give a simple proof that the Dobrushin and Shlosman's complete analyticity condition, for a Gibbs specification with finite range summable interaction, implies uniform logarithmic Sobolev inequalities. This result has been previously proven using several different techniques. The advantage of our approach is that it relies almost entirely on a general property of the entropy, while very little is assumed on the Dirichlet form. No topology is introduced on the single spin space, thus discrete and continuous spins can be treated in the same way.

Key Words: Entropy, Logarithmic Sobolev inequalities, Gibbs measures

---

Mathematics Subject Classification 2000: 82B20, 82C20, 39B62

1. Introduction

Logarithmic Sobolev inequalities have been introduced in [Gr1] where it has been shown that

$$\int_{\mathbb{R}^d} f^2(x) \log |f(x)| \, \gamma_d(dx) \leq \int_{\mathbb{R}^d} |\nabla f(x)|^2 \, \gamma_d(dx) + \|f\|^2_{L^2(\gamma_d)} \log \|f\|_{L^2(\gamma_d)} \,. \tag{1.1}$$

where γ_d is the Gaussian measure on $\mathbb{R}^d$. Inequality (1.1) can be written for an arbitrary symmetric Markov semigroup $P_t := e^{tL}$ on the probability space $(\Omega, \mathcal{F}, \mu)$ in the form

$$\mathrm{Ent}(f^2) \leq 2\, c\, \mathcal{E}(f) \tag{1.2}$$

where $\mathcal{E}$ is the Dirichlet form associated with the semigroup, and for any $g \geq 0$, $\mathrm{Ent}(g)$ stands for the entropy of g w.r.t. μ, defined as

$$\mathrm{Ent}(g) := \int_\Omega g \, \log g \, d\mu - \int_\Omega g \, d\mu \, \log \int_\Omega g \, d\mu \,.$$

Unlike classical Sobolev inequalities, (1.1) is dimension independent and remain both meaningful and valid in infinite dimensions. This fundamental feature is based on the well known factorization (or tensorization) property, expressed in the Faris' additivity theorem [Fa], which can be stated as follows: assume that (1.2) holds for two semigroups e^{-tL_i}, $i = 1, 2$, acting respectively on $(\Omega_i, \mathcal{F}_i, \mu_i)$, then it also holds, *with the same constant* c, for the semigroup e^{-tL}, with $L := L_1 \otimes I + I \otimes L_2$ acting on the product space $(\Omega_1 \times \Omega_2)$. The factorization property clearly makes the logarithmic Sobolev inequalities (LSI) a suitable instrument for infinite dimensional analysis. In particular, one field where the application of the LSI has been remarkably successful is the theory of Gibbs random fields. The main reason for this interest is the equivalence [Gr1] between (1.2) and the Nelson's hypercontractivity bound [Ne]

$$\|P_t\|_{L^q \leftarrow L^p} \leq 1 \qquad \text{if } e^{2t/c} \geq (q-1)/(p-1) \tag{1.3}$$

which, in turns, is a natural tool for studying the speed of convergence of quantities like $P_t f$ to their limit value (as t goes to ∞) $\int_\Omega f \, d\mu$. More precisely (1.3) allows one to convert L^2 convergence into a (much) stronger statement.

One fundamental problem is to find conditions under which a LSI is satisfied, for a given Gibbsian specification, *uniformly in the volume and the boundary condition*. In the trivial case of absence of interaction, the Gibbs measure is just a product of simple factors, thus, thanks to the factorization property, a uniform LSI is directly implied by the validity of a LSI for each of these factors. The problem becomes interesting when the interaction is non zero, and one is tempted to conjecture that if the interaction is weak, so that the associated Gibbs measure is "almost" a product, then the conclusion is the same as in the product case. A series of remarkable papers (see [HS1], [HS2], [SZ1], [SZ2], [SZ3], [MO1], [MO2], [LY] and reference therein) has shed conclusive light on the subject, for discrete/compact single spin space, and the result is striking: a uniform LSI is *equivalent* to the well known Dobrushin and Shlosman's *complete analyticity* condition [DoSh1], [DoSh2], [DoSh3] which "almost" characterizes the absence of phase transitions (these results have been partially extended to unbounded spin systems [Ze], [Yo1], [Yo2], [Yo3], [BH], [Le2]).

In this paper we give a new proof of the fact that complete analyticity implies a uniform LSI. The proof we present follows the general iterative strategy devised in Theorem 4.6 of [Ma], but eliminate most of its technical complications, and it is considerably easier than any existing proof. The argument relies almost completely on a general property (quasi–factorization) of the entropy, which holds in an arbitrary probability space. The advantage of this approach is that very little is assumed on the Dirichlet form and no topological hypotheses are to be made on the spin space. The treatment is thus identical for discrete and continuous spins.

At the completion of this work we learned from P. Dai Pra, A. M. Paganoni and G. Posta that they had independently obtained an inequality like to (2.10) below, in order to prove exponential decay of the entropy for a system of $\mathbb{N}$–valued spins [DPP].

Acknowledgements. The author thanks Fabio Martinelli for many helpful conversations.

2. Quasi–factorization of the entropy

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and let $\mathcal{F}_1$, $\mathcal{F}_2$ be two sub–σ–algebras of $\mathcal{F}$. We denote with $\mu(f)$ and $\mathrm{Var}(f)$ the expectation and the variance of f. If f is non negative and such that $f \log^+ f \in L^1(\mu)$, $\mathrm{Ent}(f)$ stands for the *entropy* of f w.r.t. μ, given by

$$\mathrm{Ent}(f) := \mu\Big[f \log \frac{f}{\mu(f)}\Big]\,.$$

We also let, for $i = 1, 2$, $\mu_i(f) := \mu(f \,|\, \mathcal{F}_i)$. μ_i is a linear operator from $L^1(\Omega, \mathcal{F}, \mu)$ to $L^1(\Omega, \mathcal{F}_i, \mu)$ with norm equal to 1. The restriction of μ_i to L^p is also a contraction from $L^p(\Omega, \mathcal{F}, \mu)$ to $L^p(\Omega, \mathcal{F}_i, \mu)$. For $p \in [1, \infty]$, we let $\|f\|_p := \|f\|_{L^p(\mu)}$. We define the conditional variance and the conditional entropy of f as

$$\begin{aligned} &\mathrm{Var}_i(f) := \mu_i(f^2) - \mu_i(f)^2 && f \in L^2(\mu) \\ &\mathrm{Ent}_i(f) := \mu_i[f \log f] - \mu_i(f) \log \mu_i(f) && f \geq 0,\ f \log^+ f \in L^1(\mu)\,. \end{aligned} \tag{2.1}$$

The quantities $\mathrm{Var}_i(f)$ and $\mathrm{Ent}_i(f)$ are both elements of $L^1(\mu)$ since $x \log x \geq -1/e$ and, by Jensen's inequality, $\mu_i(f) \log \mu_i(f) \leq \mu_i(f \log f)$, a.s. It is well known (see for instance Proposition 2.2 in [Le1]) that if $\mathcal{F}_1$ and $\mathcal{F}_2$ are independent we have

$$\mathrm{Var}(f) \leq \mu[\mathrm{Var}_1(f) + \mathrm{Var}_2(f)] \tag{2.2}$$

$$\mathrm{Ent}(f) \leq \mu[\mathrm{Ent}_1(f) + \mathrm{Ent}_2(f)]\,. \tag{2.3}$$

Apart from technicalities which prevent the occurrence of possible divergences, and which will be dealt with in the next section, the proof, say for the entropy, is very simple:

$$\mathrm{Ent}(f) = \mu\Big[f \log \frac{f}{\mu_1 f}\Big] + \mu\Big[f \log \frac{\mu_1 f}{\mu_2 \mu_1 f}\Big] + \mu\Big[f \log \frac{\mu_2 \mu_1 f}{\mu f}\Big] \tag{2.4}$$

$$= \mu\mu_1\Big[f \log \frac{f}{\mu_1 f}\Big] + \mu\mu_2\Big[f \log \frac{\mu_1 f}{\mu_2 \mu_1 f}\Big] + \mu\Big[f \log \frac{\mu_2 \mu_1 f}{\mu f}\Big]\,. \tag{2.5}$$

The first term of (2.5) is equal to $\mu(\mathrm{Ent}_1(f))$. The second term is less than or equal to $\mu(\mathrm{Ent}_2(f))$, since, in general one has (Proposition 2.2 in [Le1])

$$\mathrm{Ent}(f) = \sup\{\mu(fg) : \mu(e^g) \le 1\}\,.$$

Finally the last term is zero, because $\mathcal{F}_1$, $\mathcal{F}_2$ are independent, and thus $\mu_2\mu_1 f = \mu f$ a.s. Hence (2.3) follows.

It is natural to guess that inequalities (2.2), (2.3) are stable against appropriate "perturbations" of the hypothesis of independence of the σ–algebras $\mathcal{F}_1$, $\mathcal{F}_2$. The independence assumption can be stated by saying that $\mu_2(f)$ is a.s. equal to $\mu(f)$ whenever f is measurable w.r.t. $\mathcal{F}_1$. Hence one may look for a "weak dependence" condition of the form

$$|\mu_2(f) - \mu(f)| \text{ is small in some sense} \qquad \forall f \in L^1(\Omega, \mathcal{F}_1, \mu)\,.$$

In [BCC], Lemma 3.1, we have proven the following result

Proposition [BCC]. *Assume that for some* $\varepsilon \in [0, \sqrt{2} - 1)$, $p \in [1, \infty]$, *we have*

$$\begin{aligned} \|\mu_1 g - \mu g\|_p &\le \varepsilon\, \|g\|_p \qquad \forall g \in L^p(\Omega, \mathcal{F}_2, \mu) \\ \|\mu_2 g - \mu g\|_p &\le \varepsilon\, \|g\|_p \qquad \forall g \in L^p(\Omega, \mathcal{F}_1, \mu) \end{aligned} \tag{2.6}$$

Then

$$\mathrm{Var}(f) \le (1 - 2\varepsilon - \varepsilon^2)^{-1}\, \mu[\mathrm{Var}_1(f) + \mathrm{Var}_2(f)] \qquad \forall f \in L^2(\mu)\,. \tag{2.7}$$

In the following section we show that a stronger notion of "quasi–independence" than (2.6), yields an analogous quasi-factorization property for the entropy.

Proposition 2.1. *There exist* $\alpha < \infty$, $\vartheta : [0, 1) \mapsto \mathbb{R}_+$, *with* $\limsup_{\varepsilon \to 0} (\vartheta(\varepsilon)/\varepsilon) \le \alpha$, *such that the following holds: if for some* $\varepsilon \in [0, 1)$,

$$\|\mu_2(g) - \mu(g)\|_\infty \le \varepsilon \|g\|_1 \qquad \forall g \in L^1(\Omega, \mathcal{F}_1, \mu) \tag{2.8}$$

then, for all functions f *such that* $f^2 \log^+ f^2 \in L^1(\mu)$, *we have*

$$\mathrm{Ent}(f^2) \le \mu[\mathrm{Ent}_1(f^2) + \mathrm{Ent}_2(f^2)] + \vartheta(\varepsilon)\, \mathrm{Var}(f) \tag{2.9}$$

$$\mathrm{Ent}(f^2) \le \mu[\mathrm{Ent}_1(f^2) + \mathrm{Ent}_2(f^2)] + \vartheta(\varepsilon)\, \mathrm{Ent}(f^2)\,. \tag{2.10}$$

In particular one can take $\vartheta(\varepsilon) := 84\varepsilon/(1 - \varepsilon)^2$.

Remarks.

(i) When $\mathcal{F}_1$, $\mathcal{F}_2$ are independent, assumptions (2.6) and (2.8) hold with $\varepsilon = 0$, thus previous results generalize the factorization properties (2.2), (2.3).

(ii) Hypothesis (2.8) appears to be very strong but cannot be qualitatively improved in general, at least for getting (2.9). Take, in fact, $\Omega := \Omega_1 \times \Omega_2$, where $\Omega_1 = \Omega_2 = \mathbb{N}$. Choose $i, j \in \mathbb{N}$ and let $f(k, l) := \mathbb{I}_{\{k=i,\, l=j\}}$. A straightforward calculation shows that (2.9) implies that

$$e^{-\vartheta(\varepsilon)} \le \frac{\mu(\omega_1 = i \,|\, \omega_2 = j)}{\mu(\{i\} \times \Omega_2)} \le e^{\vartheta(\varepsilon)}$$

which, being valid for all $i,j \in \mathbb{N}$, in turns implies that if $g \geq 0$ is measurable w.r.t. $\mathcal{F}_1$, then

$$e^{-\vartheta(\varepsilon)}\mu(g) \leq \mu(g \,|\, \omega_2 = n) \leq e^{-\vartheta(\varepsilon)}\mu(g) \qquad \forall n \in \mathbb{N}\,.$$

A function g with no definite sign can be written as $g = g_+ - g_-$, and one gets

$$|\mu(g \,|\, \omega_2 = n) - \mu(g)| \leq (e^{\vartheta(\varepsilon)} - 1)\,\mu(|g|) \qquad \forall n \in \mathbb{N}\,.$$

(iii) In section 4 we show that the well known *complete analyticity* condition for a Gibbs specification implies (2.8) for a suitable choices of $\mathcal{F}_1$, $\mathcal{F}_2$.

3. Proof of Proposition 2.1

Step (1). Proof that (2.9) implies (2.10).
Inequality (2.9) can be applied to $|f|$ to get the (trivial) improvement

$$\mathrm{Ent}(f^2) \leq \mu[\mathrm{Ent}_1(f^2) + \mathrm{Ent}_2(f^2)] + \vartheta(\varepsilon)\,\mathrm{Var}(|f|)\,.$$

Since for a non negative function g, $\mathrm{Var}(g) \leq \mathrm{Ent}(g)$ (this is a consequence, for instance, of Lemma 1 in [LO]), (2.10) follows from (2.9).

Step (2). It is sufficient to prove (2.9) for all functions f such that $b^{-1} \leq f^2 \leq b$ for some $b > 1$.
Let in fact $f^2 \log^+ f^2 \in L^1(\mu)$ and let $g_n := (|f_n| \wedge n) \vee n^{-1}$. Then by repeatedly applying the dominated convergence theorem, one easily shows that $\mathrm{Var}(g_n)$, $\mathrm{Ent}(g_n^2)$ and $\mu[\mathrm{Ent}_i(g_n^2)]$ converge respectively to $\mathrm{Var}(|f|)$, $\mathrm{Ent}(f^2)$ and $\mu[\mathrm{Ent}_i(f^2)]$.

Step (3). If $b^{-1} \leq f^2 \leq b$ for some $b > 1$, then (2.9) holds.
For such an f we can proceed as in (2.4), (2.5) and obtain

$$\mathrm{Ent}(f^2) \leq \mu[\mathrm{Ent}_1(f^2) + \mathrm{Ent}_2(f^2)] + \mu\Big[f^2 \log \frac{\mu_2\mu_1(f^2)}{\mu(f^2)}\Big]\,. \tag{3.1}$$

Thus, in order to prove (2.9) with the explicit expression for ϑ given in the statement, it is sufficient to show that

$$\mu\Big[f^2 \log \frac{\mu_2\mu_1(f^2)}{\mu(f^2)}\Big] \leq 84\,\varepsilon(1-\varepsilon)^{-2}\;\mathrm{Var}(f)\,. \tag{3.2}$$

By assumption (2.8)

$$\|\mu_2\mu_1 f^2 - \mu f^2\|_\infty = \|\mu_2\mu_1 f^2 - \mu\mu_1 f^2\|_\infty \leq \varepsilon\|\mu_1 f^2\|_1 \leq \varepsilon\|f^2\|_1$$

i.e.

$$(1-\varepsilon)\,\mu(f^2) \leq \mu_2\mu_1(f^2) \leq (1+\varepsilon)\,\mu(f^2) \quad \mu\text{–a.s.} \tag{3.3}$$

If $\mu(f) = 0$, from (3.3) we get

$$\mu\Big[f^2 \log \frac{\mu_2\mu_1(f^2)}{\mu(f^2)}\Big] \leq \log(1+\varepsilon)\,\mu(f^2) \leq \varepsilon\,\mathrm{Var}(f) \tag{3.4}$$

and (3.2) holds with $\vartheta(\varepsilon)=\varepsilon$. In order to prove (3.2) when $\mu(f)\neq 0$, we proceed as in the proof of Rothaus inequality [Ro], given in [DeSt, p.246]. Without loss of generality we can assume $\mu(f)=1$, and we write $f=1+t_0 g$ where g is a function with zero mean and unit variance, while $t_0:=\mathrm{Var}(f)$. For $t\geq 0$ we let $f_t:=1+tg$. Since f_t^2 has no reason to satisfy a lower bound $f_t^2\geq b^{-1}$, we introduce a regularizing parameter γ, with $\gamma\in(0,\varepsilon]$ and define

$$\varphi_\gamma(t):=\mu\Big[f_t^2\log\frac{\mu_2\mu_1(f_t^2)+\gamma}{\mu(f_t^2)}\Big]=\mu\Big[f_t^2\log\frac{\mu_2\mu_1(f_t^2)+\gamma}{1+t^2}\Big]\qquad t\in\mathbb{R}.$$

We claim that, for all $\gamma\in(0,\varepsilon]$

$$\varphi_\gamma(0)\leq\gamma \tag{3.5}$$

$$\varphi_\gamma'(0)=0 \tag{3.6}$$

$$\varphi_\gamma''(t)\leq 168\,\varepsilon\,(1-\varepsilon)^{-2}\quad\forall t\geq 0 \tag{3.7}$$

Given (3.5), (3.6), (3.7) the proof of (3.2) easily follows. In fact $\varphi_\gamma(t_0)\leq\gamma+84\,\varepsilon(1-\varepsilon)^{-2}t_0^2$. Moreover, since

$$b\log(b^{-2})\leq f^2\log\frac{\mu_2\mu_1(f^2)+\gamma}{\mu(f^2)}\leq b\log[b(b+\gamma)]$$

we have, by dominated convergence,

$$\mu\Big[f^2\log\frac{\mu_2\mu_1(f^2)}{\mu(f^2)}\Big]=\lim_{n\to\infty}\varphi_{1/n}(t_0)\leq 84\,\varepsilon(1-\varepsilon)^{-2}\ \mathrm{Var}(f)\,.$$

In the rest of this section we show that (3.5), (3.6), (3.7) hold, completing the proof of the theorem. The first inequality (3.5) is trivial. Thanks to the parameter γ we can safely differentiate under the expectations, so, using $\mu(g)=0$ and $\mu(g^2)=1$, we get

$$\varphi_\gamma'(t)=2\mu\Big[f_t g\log\frac{\mu_2\mu_1(f_t^2)+\gamma}{1+t^2}\Big]+2\mu\Big[f_t^2\frac{\mu_2\mu_1(f_t g)}{\mu_2\mu_1(f_t^2)+\gamma}\Big]-2t\,. \tag{3.8}$$

Hence

$$\varphi_\gamma'(0)=2\,\log(1+\gamma)\,\mu(g)+2\,(1+\gamma)^{-1}\mu[\mu_2\mu_1(g)]=0\,.$$

Differentiating again (3.8) we obtain

$$\begin{aligned}\varphi_\gamma''(t)=&\,2\mu\Big[g^2\log\frac{\mu_2\mu_1(f_t^2)+\gamma}{1+t^2}\Big]+2\mu\Big[f_t^2\frac{\mu_2\mu_1(g^2)}{\mu_2\mu_1(f_t^2)+\gamma}\Big]\\&+8\mu\Big[gf_t\frac{\mu_2\mu_1(gf_t)}{\mu_2\mu_1(f_t^2)+\gamma}\Big]-4\mu\Big[f_t^2\frac{(\mu_2\mu_1(gf_t))^2}{(\mu_2\mu_1(f_t^2)+\gamma)^2}\Big]-4\frac{t^2}{1+t^2}-2\end{aligned} \tag{3.9}$$

In order to conclude the proof we observe that if we replace $\mu_2\mu_1$ with μ and take $\gamma=0$, the RHS of (3.9) is identically zero. By consequence we can write

$$\begin{aligned}\varphi_\gamma''(t)=&\,2\mu\Big[g^2\log\frac{\mu_2\mu_1(f_t^2)+\gamma}{1+t^2}\Big]+2\mu\Big[f_t^2\Big(\frac{\mu_2\mu_1(g^2)}{\mu_2\mu_1(f_t^2)+\gamma}-\frac{\mu(g^2)}{\mu(f_t^2)}\Big)\Big]\\&+8\mu\Big[gf_t\Big(\frac{\mu_2\mu_1(gf_t)}{\mu_2\mu_1(f_t^2)+\gamma}-\frac{\mu(gf_t)}{\mu(f_t^2)}\Big)\Big]-4\mu\Big[f_t^2\Big(\frac{(\mu_2\mu_1(gf_t))^2}{(\mu_2\mu_1(f_t^2)+\gamma)^2}-\frac{(\mu(gf_t))^2}{(\mu(f_t^2))^2}\Big)\Big]\\=&\,2\mu(g^2\Delta_0)+2\mu(f_t^2\Delta_1)+8\mu(gf_t\Delta_2)+4\mu(f_t^2\Delta_3)\end{aligned} \tag{3.10}$$

where $\Delta_0, \dots, \Delta_3$ are implicitly defined in the last equality. The estimate of φ''_γ will follow from the fact that for all functions h, $\mu_2\mu_1(h)$ is close to $\mu(f)$ in L^∞ sense. More precisely from assumption (2.8), using also $\|g\|_1 \le \|g\|_2 = 1$, it follows

$$\begin{aligned} (1-\varepsilon)\,(1+t^2) \le \mu_2\mu_1(f_t^2) \le (1+\varepsilon)\,(1+t^2) \quad \mu\text{–a.s.} \\ t-\varepsilon(1+t) \le \mu_2\mu_1(gf_t) \le t+\varepsilon(1+t) \quad \mu\text{–a.s.} \end{aligned} \tag{3.11}$$

The remaining rather detailed and straightforward computations have the main purpose of finding the explicit value for the function ϑ stated in the theorem, which is however not to be taken too seriously. From (3.10) we get

$$|\varphi''_\gamma(t)| \le 2\|\Delta_0\|_\infty + 2\|\Delta_1\|_\infty(1+t^2) + 8\|\Delta_2\|_\infty(1+t) + 4\|\Delta_3\|_\infty(1+t^2)\,. \tag{3.12}$$

We finally estimates $\Delta_0, \dots, \Delta_3$, using $\gamma \in (0, \varepsilon]$ and (3.11)

$$\|\Delta_0\|_\infty \le \max\{\log(1+\varepsilon), -\log(1-\varepsilon)\} \le \frac{\varepsilon}{1-\varepsilon}\,.$$

Next, we have that, μ–a.s.

$$\begin{aligned} \Delta_1 &\le \frac{1+\varepsilon}{(1-\varepsilon)(1+t^2)} - \frac{1}{1+t^2} = \frac{2\varepsilon}{1-\varepsilon}\frac{1}{1+t^2} \\ \Delta_1 &\ge \frac{1-\varepsilon}{(1+\varepsilon)(1+t^2)+\gamma} - \frac{1}{1+t^2} \ge -\frac{3\varepsilon}{1+t^2}\,. \end{aligned}$$

As for Δ_2, μ–a.s. we have

$$\begin{aligned} \Delta_2 &\le \frac{t+\varepsilon(1+t)}{(1-e)(1+t^2)} - \frac{t}{1+t^2} = \frac{\varepsilon}{1-\varepsilon}\,\frac{1+2t}{1+t^2} \\ \Delta_2 &\ge \frac{t-\varepsilon(1+t)}{(1+e)(1+t^2)+\gamma} - \frac{t}{1+t^2} \ge -\frac{\varepsilon}{1+2\varepsilon}\,\frac{1+3t}{1+t^2}\,. \end{aligned}$$

Finally, for Δ_3, μ–a.s. we have

$$|\Delta_3| \le |\Delta_2|\left|\frac{\mu_2\mu_1(gf_t)}{\mu_2\mu_1(f_t^2)+\gamma} + \frac{\mu(gf_t)}{\mu(f_t^2)}\right| \le |\Delta_2|\,\frac{2}{1-\varepsilon}\,\frac{1+2t}{1+t^2}\,.$$

Thus, collecting all together,

$$\varphi''_\gamma(t) \le \frac{8\varepsilon}{1-\varepsilon} + \frac{16\varepsilon}{(1-\varepsilon)^2}\sup_{t\ge 0}\frac{(1+3t)^2}{1+t^2} = \frac{8\varepsilon}{1-\varepsilon} + 10\,\frac{16\varepsilon}{(1-\varepsilon)^2} \le \frac{168\varepsilon}{(1-\varepsilon)^2}$$

and (3.7) follows $\square$

4. Log–Sobolev inequalities for completely analytical Gibbs random fields

Gibbs measures. We briefly recall the concept of Gibbs measures and refer the reader to [Ge] for a comprehensive introduction to the subject. We consider the d dimensional lattice $\mathbb{Z}^d$ with *sites* $x := (x_1, \ldots, x_d)$ and norm

$$|x| := \max_{i\in\{1,\ldots,d\}} |x_i|$$

The associated distance function is denoted by $d(\cdot,\cdot)$. The cardinality of $\Lambda \subset \mathbb{Z}^d$ is denoted by $|\Lambda|$. $\mathbb{F}$ is the set of all nonempty finite subsets of $\mathbb{Z}^d$. We define the exterior *n–boundary* as $\partial_n^+\Lambda := \{x \in \Lambda^c : d(x,\Lambda) \le n\}$, where Λ^c stands for the complement of Λ in $\mathbb{Z}^d$.

Given an arbitrary probability space $(S,\mathcal{E},\nu)$ (the *single spin* space), we introduce the *configuration space* $(\Omega,\mathcal{F}) := (S^{\mathbb{Z}^d},\mathcal{E}^{\mathbb{Z}^d})$. Sometimes we consider finite volume configuration spaces $(\Omega_\Lambda,\mathcal{F}_\Lambda) := (S^\Lambda,\mathcal{E}^\Lambda)$, for $\Lambda \in \mathbb{F}$. Given $\sigma \in \Omega$ and $\Lambda \subset \mathbb{Z}^d$ we denote by π_Λ the natural projection over Ω_Λ and write $\sigma_\Lambda := \pi_\Lambda(\sigma)$. If $U, V \subset \mathbb{Z}^d$ are disjoint, $\sigma_U\eta_V$ is the configuration on $U \cup V$ which is equal to σ on U and η on V. The action of the translations is defined on Ω as

$$\vartheta_x(\sigma)(y) := \sigma(y-x) \qquad x, y \in \mathbb{Z}^d.$$

If f is a function on Ω, Λ_f denotes the smallest subset of $\mathbb{Z}^d$ such that $f(\sigma)$ depends only on σ_{Λ_f}. f is called *local* if Λ_f is finite. The supremum norm of f is denoted by $\|f\|_u := \sup_{\omega\in\Omega}|f(\omega)|$.

In the following we consider a translation invariant, summable *interaction* J, of finite range r, *i.e.* a collection of functions $J = (J_A)_{A\in\mathbb{F}}$, such that $J_A : \Omega \mapsto \mathbb{R}$ is measurable w.r.t. $\mathcal{F}_A$, and

(H1) $J_{A+x} \circ \vartheta_x = J_A$ for all $A \in \mathbb{F}$, $x \in \mathbb{Z}^d$
(H2) $J_A = 0$ if the diameter of A is greater than r
(H3) $\sum_{A\in\mathbb{F}:\, A\ni 0} \|J_A\|_u < \infty$

The Hamiltonian $(H_\Lambda)_{\Lambda\in\mathbb{F}}$ associated with J is defined as

$$H_\Lambda : \Omega \ni \sigma \to \sum_{A\in\mathbb{F}:\, A\cap\Lambda\neq\emptyset} J_A \in \mathbb{R}.$$

Clearly $\|H_\Lambda\|_u \le |\Lambda|\|J\|$. For $\sigma,\tau \in \Omega$ we also let $H_\Lambda^\tau(\sigma) := H_\Lambda(\sigma_V\tau_{V^c})$ and τ is called the *boundary condition*. For each $\Lambda \in \mathbb{F}$, $\tau \in \Omega$ the (finite volume) Gibbs measure on $(\Omega,\mathcal{F})$, are given by

$$\mu_\Lambda^\tau(d\sigma) := \left(Z_\Lambda^\tau\right)^{-1} \exp[-H_\Lambda^\tau(\sigma)]\, \nu^\Lambda(d\sigma_\Lambda) \times \delta_{\Lambda^c,\tau}(d\sigma_{\Lambda^c}) \tag{4.1}$$

where Z_Λ^τ is the proper normalization factor called partition function, and $\delta_{\Lambda^c,\tau}$ is the probability measure on $(\Omega_{\Lambda^c},\mathcal{F}_{\Lambda^c})$ which gives unit mass to the configuration τ_{Λ^c}.

Given a measurable function f on Ω, $\mu_\Lambda^\tau f$ denotes expectation of f w.r.t. μ_Λ^τ, while, when the superscript is omitted, $\mu_\Lambda f$ stands for the function $\sigma \mapsto \mu_\Lambda^\sigma(f)$. $\mu_\Lambda f$ is measurable w.r.t. $\mathcal{F}_{\Lambda^c}$. Analogously, if $X \in \mathcal{F}$, $\mu_V(X) := \mu_V(\mathbb{1}_X)$, where $\mathbb{1}_X$ is the characteristic function on X. The set of measures (4.1) satisfies the DLR compatibility conditions

$$\mu_\Lambda(\mu_V(X)) = \mu_\Lambda(X) \qquad \forall\, X \in \mathcal{F} \qquad \forall\, V \subset \Lambda \in \mathbb{F}. \tag{4.2}$$

Since $\mu_V f$ is measurable w.r.t. $\mathcal{F}_{V^c}$ and since, trivially, for all g measurable w.r.t. $\mathcal{F}_{V^c}$ we have $\mu_V^\sigma(fg) = g(\sigma)\mu_V^\sigma(f)$, we get that (4.2) is equivalent to saying that $\mu_V f$ is a version of the conditional expectation $\mu_\Lambda^\tau(f \,|\, \mathcal{F}_{V^c})$. A probability measure μ on $(\Omega, \mathcal{F})$ is called a *Gibbs measure* for J if

$$\mu(\mu_V(X)) = \mu(X) \qquad \forall\, X \in \mathcal{F} \qquad \forall\, V \in \mathbb{F}. \tag{4.3}$$

Complete analyticity. In [DoSh1], [DoSh2], [DoSh3], Dobrushin and Shlosman introduced the powerful concept of complete analyticity of an interaction J, within the framework of finite single spin space S. They have shown how complete analytical interactions can be characterized by 12 equivalent conditions, and the associated Gibbs fields exhibit all regularity properties of the high–temperature regime. In particular complete analyticity implies that there is a unique Gibbs measure for J. Our basic assumption on the interaction J, is condition (IIId) in [DoSh3]. While its equivalence to the other 11 formulations of complete analyticity depends on the finiteness of S, and does not apply at our level of generality, this condition is nevertheless sufficient to prove our result for an arbitrary single spin space (see, however, remark (i) below).

In order to state this assumption we need a few definitions. For $V \subset \Lambda \in \mathbb{F}$, we define $\mu_{\Lambda,V}^\tau$ as the restriction of μ_Λ^τ to $\mathcal{F}_V$. A version of the Radon–Nikodym density of $\mu_{\Lambda,V}^\tau$ w.r.t. ν^V is given by

$$\rho_{\Lambda,V}^\tau(\sigma) := \left(Z_\Lambda^\tau\right)^{-1} \int_{\Omega_{\Lambda\setminus V}} \exp[-H_\Lambda^\tau(\eta_{\Lambda\setminus V}\sigma_V)]\, \nu^{\Lambda\setminus V}(d\eta) \qquad \sigma \in \Omega_V\,.$$

Our basic hypothesis on J is then the following:

Assumption (CA). *(Complete analyticity). There exist $K > 0$, $m > 0$ such that for all $V \in \mathbb{F}$, $x \in \partial_r^+ V$, $\Delta \subset V$, and for all $\sigma, \omega \in \Omega$ with $\sigma(y) = \omega(y)$, if $y \neq x$, we have*

$$\left\| \frac{\rho_{V,\Delta}^\omega}{\rho_{V,\Delta}^\sigma} - 1 \right\|_u \leq K e^{-m d(x,\Delta)} \tag{4.4}$$

Remarks. (i) Because of the sup norm in (4.4), this assumption is bound to fail for unbounded interactions, so the result we present in this section can, in practice, be mainly applied to discrete/compact spins. On the other side, logarithmic Sobolev inequalities for unbounded spins (with unbounded interaction) have been recently studied by several authors (see the references in the introduction) under the fundamental assumption of strict convexity at infinity of the Hamiltonian, which permits curvature–type arguments, like the $\Gamma_2 \geq R\Gamma$ Bakry–Emery criterion [BaEm]. For a clear analysis and review of most of these results we refer the reader to [Le2].

(ii) In [MO1] it was realized that complete analyticity is too strong a condition, and that it should be replaced by a milder assumption (called "strong mixing" in [MO1]) in which only regular volumes (*i.e.* volumes which are unions of translations of a sufficiently large given cube) are considered. Our Theorem 4.1 below can be stated using "strong mixing" rather than complete analyticity.

Logarithmic Sobolev inequalities. Given $\Lambda \in \mathbb{F}$, $\tau \in \Omega$ and a non negative function f, such that $f \log^+ f \in L^1(\mu_\Lambda^\tau)$ we define $\mathrm{Ent}_\Lambda^\tau(f)$ as the entropy of f w.r.t. μ_Λ^τ. When we write $\mathrm{Ent}_\Lambda(f)$ without the superscript τ, we mean the function $\tau \to \mathrm{Ent}_\Lambda^\tau(f)$, in analogy with $\mu_\Lambda(f)$. We consider then a "generalized" Dirichlet form $\mathcal{E}^\tau_\Lambda$. Again we define $\mathcal{E}_\Lambda(f) : \Omega \mapsto \mathbb{R}_+$ as the function $\tau \to \mathcal{E}^\tau_\Lambda(f)$. Typically [Gr2] $\mathcal{E}^\tau_\Lambda$ is the Dirichlet form associated to the generator L^τ_Λ of a symmetric, positive preserving, contraction semigroup $e^{tL^\tau_\Lambda}$. We will proceed, however, in a more abstract framework, since all we need are the following general properties of $\mathcal{E}$:

(E1) There exists a set $\mathcal{A}$ of measurable functions which is a domain for all $\{\mathcal{E}^\tau_\Lambda : \Lambda \in \mathbb{F},\ \tau \in \Omega\}$, and $\mathcal{E}^\tau_\Lambda$ maps $\mathcal{A}$ into $[0, \infty)$.

(E2) For all $V \subset \Lambda \in \mathbb{F}$, $\tau \in \Omega$, $f \in \mathcal{A}$, the function $\mathcal{E}_V(f)$ is in $L^1(\mu^\tau_\Lambda)$.

(E3) If $\Lambda = V_1 \cup V_2$, then $\mu^\tau_\Lambda[\mathcal{E}_{V_1}(f) + \mathcal{E}_{V_2}(f)] = \mathcal{E}^\tau_\Lambda(f) + \mu^\tau_\Lambda(\mathcal{E}_{V_1\cap V_2}(f))$.

We do not discuss density properties of the domain $\mathcal{A}$, and our statement will be given for functions f belonging to $\mathcal{A}$. In the specific study of a particular case, one can investigate the possibility of choosing $\mathcal{A}$ in such a way that all statements can be extended, by density, to the whole domain of $\mathcal{E}^\tau_\Lambda$.

For all $\Lambda \in \mathbb{F}$, we define the *logarithmic Sobolev constant* $c(\Lambda) \in [0, \infty]$ as the infimum of all positive real numbers c such that

$$\mathrm{Ent}^\tau_\Lambda(f^2) \le 2c\,\mathcal{E}^\tau_\Lambda(f) \qquad \forall \tau \in \Omega, \quad \forall f \in \mathcal{A}\,.$$

At this level of abstraction it is clear that we have to avoid situations where a LSI fails, for instance, already for a single spin. We need one last assumption

(E4) The quantity $c(\Lambda)$ is finite for all $\Lambda \in \mathbb{F}$.

As it will be apparent from the proof below, assumption (E4) can be replaced by the somehow milder statement that $c(\Lambda)$ is finite for all $\Lambda \in \mathbb{F}$ whose diameter does not exceed d_0, where d_0 depends on the complete analyticity constants K and m.

Having stated the hypotheses on $\mathcal{E}^\tau_\Lambda$, we mention two classical examples where they hold

(1) The first is the Glauber dynamics when $S = \{-1, 1\}$ (or any finite set). The Dirichlet form is given by $\mathcal{E}^\tau_\Lambda(f) = \mu^\tau_\Lambda(\sum_{x\in\Lambda} c_x(D_x f)^2)$, where $c_x : \Omega \mapsto \mathbb{R}_+$ are measurable, usually bounded, functions, called the transition rates, and D_x is the discrete gradient defined as follows: for $\sigma \in \Omega$, $x \in \mathbb{Z}^d$, let $\sigma^x \in \Omega$ be the configuration obtained from σ, by flipping the spin at x. Then $D_x f(\sigma) := f(\sigma^x) - f(\sigma)$. $\mathcal{A}$ can be taken as the set of all local functions.

(2) In the second example S is a connected, compact Riemannian manifold, and ν is the normalized Riemannian measure. Consider the diffusion on S^Λ, under boundary condition τ, with generator $\sum_{x\in\Lambda}(\Delta_x - \nabla_x H^\tau_\Lambda)$, where ∇_x and Δ_x are respectively the gradient and the Laplacian on the x^{th} copy of S. In this case we have $\mathcal{E}^\tau_\Lambda(f) = \mu^\tau_\Lambda(\sum_{x\in\Lambda} |\nabla_x f|^2)$ and $\mathcal{A}$ can be chosen as the set of all functions f which can be written as $f = g \circ \pi_V$ for some $V \in \mathbb{F}$, $g \in C^\infty(S^V)$.

The only nontrivial property we have to check in these cases is $(E4)$: when $J = 0$, we have $\mu^\tau_\Lambda = \nu^\Lambda$, so it is sufficient to show that a LSI holds for ν with some log-Sobolev constant $\hat{c}$. This fact is trivial in the discrete case, where the optimal constant is known (see [Le1], Ch.5), and true in case (2) (see [Wa] for upper bounds on $\hat{c}$). In order to deal with a non zero interaction, we observe that μ^τ_Λ has density $\exp(-H^\tau_\Lambda)/Z^\tau_\Lambda$ w.r.t. ν^Λ. It is well known (see Lemma 3.5 in [SZ1]) that since $\|H^\tau_\Lambda\|_u \le |\Lambda|\|J\|$, we can control $c(\Lambda)$ in terms of $\hat{c}$. More precisely we have $c(\Lambda) \le \hat{c}\, e^{4|\Lambda|\|J\|}$.

The main result in this section is

Theorem 4.1. *Let J be a translation invariant, summable interaction of finite range r such that assumptions (CA) holds, and let $\{\mathcal{E}^\tau_\Lambda : \Lambda \in \mathbb{F},\ \tau \in \Omega\}$ satisfy conditions (E1), ..., (E4). Then*

$$\sup_{\Lambda \in \mathbb{F}} c(\Lambda) < +\infty\,.$$

Proof. We start with the following result

Lemma 4.2. *Let $\Lambda \in \mathbb{F}$, and let V_1, V_2 be two subsets of Λ, such that $\Lambda = V_1 \cup V_2$. Let $l := d(\Lambda \setminus V_1, \Lambda \setminus V_2)$. Assume that*

$$|(\partial^+_r V_2) \cap \Lambda|\, K\, e^{-ml} \le 1\,. \tag{4.5}$$

Then there exists $l_0 = l_0(K, m)$ such that for all $l \ge l_0$, for all $\tau \in \Omega$

$$\operatorname{Ent}^\tau_\Lambda(f^2) \le 2\left(1 + K' e^{-ml}\right)[c(V_1) \vee c(V_2)]\left[\,\mathcal{E}^\tau_\Lambda(f) + \mu^\tau_\Lambda(\mathcal{E}_{V_1 \cap V_2}(f))\,\right] \tag{4.6}$$

where $K' = 2e\alpha K$ and α is given in Proposition 2.1.

Proof of Lemma 4.2. Let $\tau \in \Omega$ and let $F_{\Lambda,\tau} := \{\sigma \in \Omega : \sigma_{\Lambda^c} = \tau_{\Lambda^c}\}$. By (4.2) $\mu^\tau_{\Lambda,\Lambda\setminus V_1}$ is a convex combination of $\{\mu^\sigma_{V_2,\Lambda\setminus V_1} : \sigma \in F_{\Lambda,\tau}\}$, thus

$$\left\| \frac{\rho^\omega_{V_2,\Lambda\setminus V_1}}{\rho^\tau_{\Lambda,\Lambda\setminus V_1}} - 1 \right\|_u \le \sup_{\sigma,\eta \in F_{\Lambda,\tau}} \left\| \frac{\rho^\sigma_{V_2,\Lambda\setminus V_1}}{\rho^\eta_{V_2,\Lambda\setminus V_1}} - 1 \right\|_u \qquad \forall \omega \in F_{\Lambda,\tau}\,. \tag{4.7}$$

Moreover, thanks to the finite range assumption (H2), we can assume that the configurations σ, η in the RHS of (4.7) also agree on $(\partial^+_r V_2)^c$. Thus there exists a finite sequence of configurations which interpolate between σ and η, *i.e.* a sequence $(\omega_i)^n_{1=1}$, with $\omega_i \in F_{\Lambda,\tau}$, $\omega_1 = \sigma$, $\omega_n = \eta$, $n \le |(\partial^+_r V_2) \cap \Lambda|$, and such that ω_i, ω_{i+1} differ at exactly one site of $(\partial^+_r V_2) \cap \Lambda$. Thanks to assumption (CA) we obtain

$$\left(1 - K\, e^{-md(\Lambda\setminus V_2, \Lambda\setminus V_1)}\right)^n \le \frac{\rho^\sigma_{V_2,\Lambda\setminus V_1}(\xi)}{\rho^\eta_{V_2,\Lambda\setminus V_1}(\xi)} = \prod_{i=2}^n \frac{\rho^{\omega_{i-1}}_{V_2,\Lambda\setminus V_1}(\xi)}{\rho^{\omega_i}_{V_2,\Lambda\setminus V_1}(\xi)} \le \left(1 + K\, e^{-md(\Lambda\setminus V_2, \Lambda\setminus V_1)}\right)^n \tag{4.8}$$

But if $|x| \le 1/n$, we have $|(1+x)^n - 1| \le enx$, hence if (4.5) holds, from (4.7), (4.8) we get

$$\left\| \frac{\rho^\omega_{V_2,\Lambda\setminus V_1}}{\rho^\tau_{\Lambda,\Lambda\setminus V_1}} - 1 \right\|_u \le K\, e^{-md(\Lambda\setminus V_2, \Lambda\setminus V_1)+1} = K\, e^{-ml+1} \qquad \forall \omega \in F_{\Lambda,\tau}\,.$$

By consequence

$$\|\mu_{V_2}(g) - \mu^\tau_\Lambda(g)\|_{L^\infty(\mu^\tau_\Lambda)} \le K\, e^{-ml+1}\, \|g\|_{L^1(\mu^\tau_\Lambda)} \qquad \forall g \in L^1(\Omega, \mathcal{F}_{\Lambda\setminus V_1}, \mu)\,. \tag{4.9}$$

From what we said after (4.2) we can identify $\mu_{V_2}(g)$ with $\mu^\tau_\Lambda(g \,|\, \mathcal{F}_{\Lambda\setminus V_2})$. Thus, by Proposition 2.1, there is $l_0 = l_0(K, m)$ such that, for all $l \ge l_0$

$$\operatorname{Ent}^\tau_\Lambda(f^2) \le \left(1 + K' e^{-ml}\right) \mu^\tau_\Lambda\left[\,\operatorname{Ent}_{V_1}(f^2) + \operatorname{Ent}_{V_2}(f^2)\,\right]\,. \tag{4.10}$$

Finally from the definition of $c(V)$ and from assumption $(E3)$, it follows

$$\mu_\Lambda^\tau \left[\mathrm{Ent}_{V_1}(f^2) + \mathrm{Ent}_{V_2}(f^2) \right] \leq 2 \left[c(V_1) \vee c(V_2) \right] \left[\mathcal{E}_\Lambda^\tau(f) + \mu_\Lambda^\tau(\mathcal{E}_{V_1 \cap V_2}(f)) \right] \quad \square \tag{4.11}$$

Proof of Theorem 4.1. Previous lemma suggests an iterative procedure to estimate the logarithmic Sobolev constant $c(\Lambda)$, that is divide Λ roughly into two "halves" V_1, V_2, in such a way that $\Lambda = V_1 \cup V_2$ and V_1 and V_2 have an intersection "thick" enough so that (4.5) holds. Then, by (4.6) we "almost" have $c(\Lambda) \leq (1 + K' e^{-ml}) \, (c(V_1) \vee c(V_2))$. The "almost" comes of course from the extra term $\mu_\Lambda^\tau(\mathcal{E}_{V_1 \cap V_2}(f))$. A trivial upper bound for this term is $\mathcal{E}_\Lambda^\tau$, but this is fatal to the argument since it yields $c(\Lambda) \leq 2(1 + K' e^{-ml}) \, (c(V_1) \vee c(V_2))$. However it was observed in [Ma] that one can write many, say r, different replicas of inequality (4.6), each corresponding to a different choice of V_1, V_2 such that the sets $V_1 \cap V_2$ are disjoint for different replicas. At this point we can add together all the inequalities obtained and the *sum* of all the r extra terms is still bounded by $\mathcal{E}_\Lambda^\tau(f)$. In this way we get

$$c(\Lambda) \leq (1 + K' e^{-ml}) \, (1 + 1/r) \, (c(V_1) \vee c(V_2)) \, . \tag{4.12}$$

and, if r is a function of the size of Λ which goes to 0 fast enough, a chance to obtain a convergent iteration from (4.12) becomes apparent.

The actual proof requires a simple geometric construction which was already used in [BCC] for obtaining a uniform lower bound for the spectral gap of a continuous gas. We include the details below for completeness. Let $l_k := (3/2)^{k/d}$, and let $\mathbb{F}_k$ be the set of all $A \in \mathbb{F}$ which, modulo translations and permutations of the coordinates, are contained in

$$\left([0, l_{k+1}] \times [0, l_{k+2}] \times \cdots \times [0, l_{k+d}] \right) \cap \mathbb{Z}^d$$

Let also $G_k := \sup_{V \in \mathbb{F}_k} c(V)$. The idea behind this construction is that each rectangle in $\mathbb{F}_k \backslash \mathbb{F}_{k-1}$ can be obtained as a "slightly overlapping union" of two rectangles in $\mathbb{F}_{k-1}$. More precisely we have:

Proposition 4.3. *For all $k \in \mathbb{Z}_+$, for all $\Lambda \in \mathbb{F}_k \backslash \mathbb{F}_{k-1}$ there exists a finite sequence $\{V_1^{(i)}, V_2^{(i)}\}_{i=1}^{s_k}$, where $s_k := \lfloor l_k^{1/3} \rfloor$, such that, letting $\delta_k := \frac{1}{8}\sqrt{l_k} - 2$,*

(1) $\Lambda = V_1^{(i)} \cup V_2^{(i)}$ and $V_1^{(i)}, V_2^{(i)} \in \mathbb{F}_{k-1}$, for all $i = 1, \ldots, s_k$

(2) $d(\Lambda \backslash V_1^{(i)}, \Lambda \backslash V_2^{(i)}) \geq \delta_k$, for all $i = 1, \ldots, s_k$

(3) $V_1^{(i)} \cap V_2^{(i)} \cap V_1^{(j)} \cap V_2^{(j)} = \emptyset$, if $i \neq j$

Proof. Since $\Lambda \in \mathbb{F}_k$ we can assume that $\Lambda \subset ([0, b_1] \times \cdots \times [0, b_d]) \cap \mathbb{Z}^d$ with $b_n \leq l_{k+n}$, for $n = 1, \ldots, d$. Define

$$V_1^{(i)} := \left([0, b_1] \times \cdots \times [0, b_{d-1}] \times \left[0, \frac{b_d}{2} + \frac{2i}{8} \sqrt{l_k} \right] \right) \cap \Lambda$$

$$V_2^{(i)} := \left([0, b_1] \times \cdots \times [0, b_{d-1}] \times \left[\frac{b_d}{2} + \frac{2i - 1}{8} \sqrt{l_k}, \, b_d \right] \right) \cap \Lambda$$

It is straightforward to check that $V_1^{(i)}$ and $V_2^{(i)}$ belong to $\mathbb{F}_{k-1}$. In fact we know that $b_d \le l_{k+d}$, thus, for all $i = 1, \ldots, s_k$

$$\frac{b_d}{2} + \frac{2i}{8}\sqrt{l_k} \le \frac{b_d}{2} + \frac{2s_k}{8}\sqrt{l_k} \le \frac{l_{k+d}}{2} + \frac{1}{4}l_k^{5/6} = \frac{3l_k}{4} + \frac{1}{4}l_k^{5/6} \le l_k \tag{4.13}$$

which, together with the inequalities

$$b_1 \le l_{k+1}, \quad \ldots, \quad b_{d-1} \le l_{k-1+d}$$

implies that $V_1^{(i)} \in \mathbb{F}_{k-1}$. Since $V_2^{(i)}$ is smaller than $V_1^{(i)}$, it also belongs to $\mathbb{F}_{k-1}$. $V_1^{(i)}$ and $V_2^{(i)}$ are nonempty, since, using (4.13), it easy to see that, otherwise, Λ itself would belong to $\mathbb{F}_{k-1}$ which is excluded by hypothesis. We have then $d(\Lambda \backslash V_1^{(i)}, \Lambda \backslash V_2^{(i)}) \ge \frac{1}{8}\sqrt{l_k} - 2 = \delta_k$. $\square$

We can conclude the proof of Theorem 4.1. Choose a positive integer $k_0 = k_0(K, m)$ large enough such that

(i) $\delta_k \ge l_0(K, m)$ (l_0 was defined in Lemma 4.2)
(ii) $(l_{k_0+d} + 1)^d K e^{-m\delta_k} \le 1$

Let then $k \ge k_0$, $\Lambda \in \mathbb{F}_k$, and let $\{V_1^{(i)}, V_2^{(i)}\}_{i=1}^{s_k}$ be the sequence given in Proposition 4.3. Properties (i) and (ii) allow us to apply Lemma 4.2 and obtain, for $i = 1, \ldots, s_k$

$$\mathrm{Ent}_\Lambda^\tau(f^2) \le 2\left(1 + K' e^{-m\delta_k}\right)[c(V_1^{(i)}) \vee c(V_2^{(i)})]\left[\mathcal{E}_\Lambda^\tau(f) + \mu_\Lambda^\tau(\mathcal{E}_{V_1^{(i)} \cap V_2^{(i)}}(f))\right]. \tag{4.14}$$

Thanks to (3) of Proposition 4.3, $\sum_{i=1}^{s_k} \mu_\Lambda^\tau(\mathcal{E}_{V_1^{(i)} \cap V_2^{(i)}}(f)) \le \mathcal{E}_\Lambda^\tau(f)$, so when we sum (4.14) for $i = 1, \ldots, s_k$, and divide by s_k we get

$$\mathrm{Ent}_\Lambda^\tau(f^2) \le 2G_{k-1}\left(1 + K' e^{-m\delta_k}\right)\left[1 + \frac{1}{s_k}\right]\mathcal{E}_\Lambda^\tau(f) \tag{4.15}$$

which yields

$$G_k \le G_{k-1}\left(1 + K' e^{-m\delta_k}\right)\left[1 + \frac{1}{s_k}\right] \qquad \forall k \ge k_0\,. \tag{4.16}$$

From the iteration of (4.16) we obtain $G_k \le M G_{k_0}$, for all $k \ge k_0$, where

$$M := \prod_{k=k_0}^{\infty}\left\{\left(1 + K' e^{-m\delta_k}\right)\left[1 + \frac{1}{s_k}\right]\right\} < \infty$$

thanks to the explicit expressions of δ_k and s_k. Finally assumption (E4) guarantees that G_{k_0} is finite, and the theorem follows $\square$

References

[BaEm] D. Bakry and M. Emery, *Diffusions hypercontractives*, Séminaire de probabilités, XIX. Springer, Berlin, 1985, pp. 177–206.

[BCC] L. Bertini, N. Cancrini and F. Cesi, *The spectral gap for a Glauber–type dynamics in a continuous gas*, Preprint (2000)

[BH] T. Bodineau and B. Helffer, *The log-Sobolev inequality for unbounded spin systems*, J. Funct. Anal. **166**, no. 1, 168–178 (1999)

[DeSt] J.D. Deuschel and D.W. Stroock, *Large deviations*, Academic Press Inc., Boston, MA, 1989

[DoSh1] R. L. Dobrushin and S. B. Shlosman, *Constructive criterion for the uniqueness of Gibbs field*, Statistical physics and dynamical systems (Jaffe Fritz and Szász, eds.), Birkhäuser Boston, Mass., 1985, pp. 347–370

[DoSh2] R. L. Dobrushin and S. B. Shlosman, *Completely analytical Gibbs fields*, Statistical physics and dynamical systems (Jaffe Fritz and Szász, eds.), Birkhäuser Boston, Mass., 1985, pp. 371–403

[DoSh3] R. L. Dobrushin and S. B. Shlosman, *Completely analytical interactions: constructive description*, J. Statist. Phys. **46**, no. 5-6, 983–1014 (1987)

[DPP] P. Dai Pra, A. M. Paganoni and G. Posta, *Entropy inequalities for unbounded spin systems*, in preparation

[Fa] W. G. Faris, *Product spaces and Nelson's inequality*, Helv. Phys. Acta **48**, no. 5/6, 721–730 (1975)

[Ge] H. O. Georgii, *Gibbs measures and phase transitions*, Walter de Gruyter, 1988.

[Gr1] L. Gross, *Logarithmic Sobolev inequalities*, Am. J. Math. **97** 1061–1083 (1976)

[Gr2] L. Gross, *Logarithmic Sobolev inequalities and contractivity properties of semigroups*, Dirichlet forms. Lect. Notes in Math., vol. 1563, Springer–Verlag, 1993, pp. 54–88.

[HS1] R. Holley and D. Stroock, *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Statist. Phys. **46**, no. 5-6, 1159–1194 (1987)

[HS2] R. A. Holley and D. W. Stroock, *Uniform and L^2 convergence in one-dimensional stochastic Ising models*, Comm. Math. Phys. **123**, no. 1, 85–93 (1989)

[Le1] M. Ledoux, *Concentration of measure and logarithmic Sobolev inequalities*, Séminaire de Probabilités. XXXIII. Lecture Notes in Mathematics 1709. Springer, Berlin, 1999, pp. 120–216.

[Le2] M. Ledoux, *Logarithmic Sobolev inequalities for unbounded spin systems revisited*, Preprint (1999)

[LO] R. Latała and K. Oleszkiewicz, *Between Sobolev and Poincaré*, Preprint (2000)

[LY] S. L. Lu and H. T. Yau, *Spectral Gap and Logarithmic Sobolev Inequality for Kawasaki and Glauber Dynamics*, Commun. Math. Phys. **156**, 399–433 (1993)

[Ma] F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, Lectures on probability theory and statistics (Saint-Flour, 1997). Lecture Notes in Mathematics 1717 (Berlin), Springer, Berlin, 1999, pp. 93–191.

[MO1] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region I: The attractive case*, Commun. Math. Phys. **161**, 447–486 (1994)

[MO2] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region II: The general case*, Commun. Math. Phys. **161**, 487–514 (1994)

[Ne] E. Nelson, *The free Markoff field*, J. Functional Analysis **12**, 211–227 (1973)

[Ro] O. Rothaus, *Analytic inequalities, isoperimetric inequalities, and logarithmic Sobolev inequalities*, J. Funct. Anal. **64**, 296–313 (1985)

[SZ1] D.W. Stroock and B. Zegarlinski, *The logarithmic Sobolev inequality for continuous spin systems on a lattice*, J. Funct. Anal. **104**, no. 2, 299–326 (1992)

[SZ1] D. W. Stroock and B. Zegarlinski, *The equivalence of the logarithmic Sobolev inequality and the Dobrushin–Shlosman mixing condition*, Commun. Math. Phys. **144**, 303–323 (1992)

[SZ3] D. W. Stroock and B. Zegarlinski, *The logarithmic Sobolev inequality for discrete spin on a lattice*, Commun. Math. Phys. **149**, 175 (1992)

[Wa] F.-Y. Wang, *On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups*, Probab. Theory Related Fields **108**, no. 1, 87–101 (1997)

[Yo1] N. Yoshida, *The log-Sobolev inequality for weakly coupled lattice field*, Probab. Theory Related Fields **115**, 1–40 (1999)

[Yo2] N. Yoshida, *Application of log-Sobolev inequality to the stochastic dynamics of unbounded spin systems on the lattice*, J. Funct. Anal. **173**, 74–102 (2000)

[Yo3] N. Yoshida, *The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice*, Preprint (1998)

[Ze] B. Zegarlinski, *The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice*, Commun. Math. Phys. **175**, 401–432 (1996)