A new criterion for the logarithmic Sobolev inequality

Felix Otto and Maria G. Reznikoff [a,b]

[a] *Inst. für Angewandte Mathematik, Universität Bonn, Germany, otto@iam.uni-bonn.de*

[b] *Mathematics Dept., Princeton University, Princeton, USA, mrezniko@math.princeton.edu*

Abstract

We present a criterion for the logarithmic Sobolev inequality (LSI) on the product space $X_1 \times \ldots \times X_N$. We have in mind an N–site lattice, unbounded continuous spin variables, and Glauber dynamics. The interactions are described by the Hamiltonian H of the Gibbs measure. The criterion for LSI is formulated in terms of the LSI constants of the single–site conditional measures and the size of the off–diagonal entries of the Hessian of H. It is optimal for Gaussians with positive covariance matrix. To illustrate, we give two applications: one with weak interactions and one with order–one interactions and a decay of correlations condition.

Key words: Logarithmic Sobolev inequality, decay of correlations, Glauber dynamics
60K35 (primary), 60J25, 82B21

1 Introduction

The logarithmic Sobolev inequality (LSI) was introduced by Gross [G]. It is attached to a Markov semi–group P_t with reversible invariant measure μ. We refer to the recent survey paper [GZ, Chapters 1 - 4] for a general introduction within the framework of Γ_1–calculus. Like the spectral gap inequality (SGI), which is analytically speaking a Poincaré inequality for the measure μ, LSI yields exponential convergence of the Markov semi–group to equilibrium with a rate given by the constant in the inequality. Like the classical Sobolev–Poincaré inequalities, the LSI also expresses an improved integrability. Gross made two important observations: On the one hand, this improved integrability is strong enough to yield hypercontractivity for the Markov semi–group P_t; see for instance [GZ, Theorem 4.1]. Hypercontractivity is a sharpened statement of the trend to equilibrium; see [GZ, Section 4.1]. On the other hand, the improved integrability is weak enough (the gain is just a logarithm) to be stable under Cartesian products (of the Markov semi–groups and their reversible invariant measures); see [GZ, Theorem 4.4] and Remark 1. These are the features that make the LSI suitable for spin systems.

There are only a few sufficient criteria for LSI. The first important criterion, due to Holley & Stroock [HS], is perturbative in nature and not ideally suited for spin systems. The second important criterion, due to Bakry & Emery [BE], is non–perturbative, but structurally quite restrictive, cf. Remark 2. The criterion of Bakry & Emery is based on the Γ_2–calculus, which essentially requires a Riemannian spin space. This is also the framework we adopt. We shall frequently refer to [L] for a nice review.

Our main result (Theorem 1) is a clean sufficient criterion for LSI. We consider a Gibbs measure μ on a product space $X_1 \times \ldots \times X_N$. We formulate the condition in terms of the Hamiltonian and the LSI constants of the single–site conditional measures. The result can be viewed as an adaptation of the above–mentioned product argument to allow for coupling, cf. Remark 1. It is indeed important to start from the product argument, since a naive application of the Holley–Stroock principle (see for instance [L, Lemma 1.2]) would yield an LSI–constant that increases exponentially with the number N of sites, no matter how weak the interaction. We require weaker hypotheses than the Bakry-Emery principle (cf. Remark 2): We do not require strict convexity of the Hamiltonian. Moreover, for $X_i = \mathbb{R}$ and attractive interactions, the bound of Theorem 1 is sharp for Gaussians (cf. Remark 4). For the SGI of a Gibbs measure, a result similar to Theorem 1 and somewhat stronger is proved by Ledoux [L], cf. Remark 3.

Earlier work of Royer [R, Théorème 5.2.1] based on Zegarlinski's iterative method produces a similar, but weaker, bound for the LSI constant (cf. Remark 5). Introduced by Zegarlinski in 1990 [Z1], the iterative technique was applied and developed by Zegarlinski [Z2,Z3] and Stroock & Zegarlinski [SZ].

A second approach that has been widely used in the analysis of spin systems is the Lu–Yau martingale method (introduced in [LY] and also reviewed in [L, Section 5]). The martingale method relies on the LSI for *marginals* (i. e. averaged out versions) of the conditional measures; see for instance [LY, (6.3)], [Yo, Lemma 3.2], [L, Proposition 4.1]. In the case of unbounded spin space, these LSI's for marginals rely in turn on global spectral gap estimates, cf. [Yo, Theorem 2.2] and [L, Proposition 3.1]. Recent progress by Blower and Bolley

[BB, Theorem 1.3] also transforms information about LSI for conditional and marginal measures into a global LSI.

Our functional analytic approach combines the advantages of both approaches described above: It avoids the fixed point iteration and it requires the LSI for conditional measures, but not marginals. It grows out of work presented in [O,OV,GORV]. See Section 3 for additional comments regarding connections among the different methods.

We begin in Subsection 1.1 by presenting the main result. In Subsection 1.2 we state a two–scale criterion for LSI (introduced in [GORV]) which we will use in the applications. Section 2 contains the examples. In Section 3 we present some auxiliary results, and finally in Section 4 we give the proofs.

1.1 Main result and discussion

We deal only with Euclidean spaces X, although our arguments would also go through for Riemannian manifolds. Norms $|\cdot|$ and the notion of gradient ∇ are derived from the Euclidean structure.

The logarithmic Sobolev inequality can be defined in the following way:

Definition 1 (LSI) *Let $\Phi(x) := x\log x$. The probability measure $\mu(dx)$ on X satisfies the logarithmic Sobolev inequality LSI(ρ) with constant ρ if*

$$\forall\, f(x)\geq 0 \quad \int \Phi(f)\,d\mu - \Phi\left(\int f\,d\mu\right) \;\leq\; \frac{1}{\rho}\int \frac{1}{2f}|\nabla f|^2\,d\mu. \tag{1}$$

We recall the disintegration of a probability measure into the conditional probability measure $\mu(dx_1|x_2)$ and the corresponding marginal $\bar{\mu}(dx_2)$:

Definition 2 (Conditional and marginal measures) *To any probability measure $\mu(dx_1dx_2)$ on $X_1 \times X_2$ we associate the marginal $\bar{\mu}(dx_1)$, a probability measure on X_1, and the family of conditional measures $\mu(dx_2|x_1)$, probability measures on X_2, via*

$$\forall\, \zeta(x_1,x_2) \qquad \int_{X_1\times X_2} \zeta\, d\mu \;=\; \int_{X_1}\int_{X_2} \zeta(x_1,x_2)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1).$$

Our main result is:

Theorem 1 *Let $X_1,\cdots,X_N$ be Euclidean spaces and $\mu(dx_1\cdots dx_N)$ a probability measure on the product space $X_1\times\cdots\times X_N$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

We assume that for all $i<j\in\{1,\cdots,N\}$ there exists $\kappa_{ij}<\infty$ such that the Hamiltonian $H(x_1,\cdots,x_N)=-\ln\frac{d\mu}{d\mathcal{L}}$ satisfies

$$\forall\,(x_1,\cdots,x_N) \quad |\nabla_i\nabla_j H(x_1,\cdots,x_N)| \;\le\; \kappa_{ij}. \tag{2}$$

Here and in what follows, $|\nabla_i\nabla_j H(x_1,\cdots,x_N)|$ denotes the operator norm of the bilinear form $\nabla_i\nabla_j H(x_1,\cdots,x_N)$ on $X_i\times X_j$.

We assume that for all $i\in\{1,\cdots,N\}$ there exists $\rho_i>0$ such that

$$\begin{aligned}&\forall\,(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N)\\ &\mu(dx_i|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N) \quad \text{satisfies } LSI(\rho_i).\end{aligned} \tag{3}$$

Consider the symmetric $N \times N$–matrix A defined by

$$\begin{aligned} A_{ij} &= -\kappa_{ij} \text{ for } i < j \in \{1, \cdots, N\}, \\ A_{ii} &= \rho_i \quad \text{for} \quad i \in \{1, \cdots, N\}. \end{aligned} \tag{4}$$

We assume that there exists $\rho > 0$ such that we have in the sense of quadratic forms

$$A \ \geq \ \rho \, \mathrm{id}. \tag{5}$$

Then

$$\mu(dx_1 \cdots dx_N) \quad \text{satisfies LSI}(\rho). \tag{6}$$

Remark 1 (Product measures) *Assume that we are in the setting of Theorem 1 but that in addition μ is a product measure, i. e.*

$$\mu(dx_1 \cdots dx_N) \ = \ \mu_1(dx_1) \ldots \mu_N(dx_N). \tag{7}$$

It has been known since its origins [G, Remark 3.3] (see [L, Lemma 1.1] for a modern presentation) that the LSI is compatible with taking products. More precisely, if for each $i \in \{1, \cdots, N\}$, there exists $\rho_i > 0$ such that

$$\mu_i(dx_i) \quad \text{satisfies LSI}(\rho_i),$$

then

$$\mu(dx_1 \ldots dx_N) \quad \text{satisfies LSI}(\rho)$$

with

$$\rho \ := \ \min\{\rho_1, \cdots, \rho_N\}. \tag{8}$$

Theorem 1 matches this bound for product measures: In case of (7), we have $\nabla_i \nabla_j H \equiv 0$ for all $i < j \in \{1, \cdots, N\}$. Therefore, we may choose $\kappa_{ij} = 0$, to *the effect that the optimal ρ in (5) is precisely (8). Theorem 1 can be interpreted as a perturbation of the above product property.*

Remark 2 (The criterion of Bakry–Emery) *In this remark, we relate Theorem 1 to the Bakry–Emery criterion [BE]; see for instance [L, Corollary 1.6] for an efficient proof. In the notation of Theorem 1, the Bakry–Emery principle reads as follows: Consider the symmetric $N \times N$–matrix $A(x_1, \cdots, x_N)$ defined by*

$$A_{ij}(x_1, \cdots, x_N) \;=\; \nabla_i \nabla_j H(x_1, \cdots, x_N) \quad \textit{for } \; i, j \in \{1, \cdots, N\}.$$

If $\rho > 0$ is such that

$$\forall\,(x_1, \cdots, x_N) \quad A(x_1, \cdots, x_N) \;\geq\; \rho\,\mathrm{id} \tag{9}$$

in the sense of quadratic forms, then

$$\mu \quad \textit{satisfies LSI}(\rho).$$

On the one hand, Theorem 1 is stronger than the Bakry–Emery principle, since it assumes less about the single site conditional measures,

$$\mu(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N).$$

On the other hand, Theorem 1 is somewhat weaker, since for $i < j \in \{1, \cdots, N\}$, the criterion (5) appeals to $A_{ij} = -\sup_{(x_1, \cdots, x_N)} |\nabla_i \nabla_j H(x_1, \cdots, x_N)|$, whereas the criterion (9) involves just $\nabla_i \nabla_j H(x_1, \cdots, x_N)$.

Remark 3 (Spectral gap inequality) *In this remark, we relate Theorem 1 to what is known about the spectral gap. Recall that a probability measure $\mu(dx)$ is said to satisfy the spectral gap estimate SGI(ρ) with constant $\rho > 0$*

provided

$$\forall\ h(x) \quad \int h^2\,d\mu - \Big(\int h\,d\mu\Big)^2 \ \le\ \frac{1}{\rho}\int |\nabla h|^2\,d\mu. \tag{10}$$

It is well–known that SGI(ρ) is a consequence of LSI(ρ), as can be seen by using $f = 1+\epsilon h$ in (1) and expanding to second order in ϵ.

In a situation analogous to Theorem 1, a somewhat stronger result is known on the level of spectral gap, cf. [L, Proposition 3.1]. We now state this result in the notation of Theorem 1: One assumes that for $i \in \{1,\cdots,N\}$, there exists $\rho_i > 0$ such that

$$\begin{array}{l}\forall\ (x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N)\\ \mu(dx_i|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N) \quad \textit{satisfies SGI}(\rho_i).\end{array}$$

One considers the symmetric $N\times N$–matrix $A(x_1,\cdots,x_N)$ defined through

$$\begin{array}{lcl} A_{ij}(x_1,\cdots,x_N) = \nabla_i\nabla_j H(x_1,\cdots,x_N) & \textit{for} & i<j \in \{1,\cdots,N\}, \\ A_{ii}(x_1,\cdots,x_N) = \qquad \rho_i & \textit{for} & i \in \{1,\cdots,N\} \end{array}$$

and assumes that there exists $\rho > 0$ such that in the sense of quadratic forms

$$\forall\ (x_1,\cdots,x_N) \quad A(x_1,\cdots,x_N)\ \ge\ \rho\,\mathrm{id}.$$

Then one has

$$\mu(dx_1\cdots dx_N) \quad \textit{satisfies SGI}(\rho).$$

Remark 4 (Gaussians) *We assume $X_i = \mathbb{R}$ for $i \in \{1,\cdots,N\}$ and let A be a symmetric positive definite $N\times N$–matrix. In this remark, we argue that Theorem 1 is optimal for Gaussians, i.e. for Hamiltonians of the form:*

$$H \ =\ -\ln\frac{d\mu}{d\mathcal{L}} \ =\ \frac{1}{2}\sum_{i,j\in\{1,\cdots,N\}} x_i\,A_{ij}\,x_j + \sum_{i\in\{1,\cdots,N\}} b_i\,x_i, \tag{11}$$

provided the coupling is attractive:

$$A_{ij} \;\leq\; 0 \quad \text{for}\, i<j\in\{1,\cdots,N\}. \tag{12}$$

Recall that the covariance matrix is given by the inverse A^{-1} of A:

$$(A^{-1})_{ij} \;=\; \int x_i\,x_j\,d\mu - \int x_i\,d\mu \int x_j\,d\mu. \tag{13}$$

Incidentally, according to Lemma 9 below, the attractive coupling (12) implies non–negative covariances

$$\int x_i\,x_j\,d\mu - \int x_i\,d\mu \int x_j\,d\mu \;\geq\; 0 \quad \text{for}\, i,j\in\{1,\cdots,N\}.$$

We also recall that for $\mu(dx)$ of the form (11) and any number $\rho>0$, we have the equivalences

$$A \;\geq\; \rho\,\mathrm{id} \quad \text{as quadratic forms} \tag{14}$$
$$\Longleftrightarrow \mu \quad \text{satisfies LSI}(\rho) \tag{15}$$
$$\Longleftrightarrow \mu \quad \text{satisfies SGI}(\rho). \tag{16}$$

Indeed, for (14) $\Longrightarrow$ (15) we refer to Remark 2 and for (15) $\Longrightarrow$ (16) to Remark 3. That (16) implies (14) can be seen as follows: For arbitrary $\xi\in\mathbb{R}^N$, choose $h(x)=\xi\cdot x$ in (10). Using (13), (10) turns into $\xi\cdot A^{-1}\xi\leq\frac{1}{\rho}|\xi|^2$, which amounts to (14).

We now give the argument for optimality: Let μ satisfy LSI(ρ). By (15) $\Longrightarrow$ (14) we must have

$$A \;\geq\; \rho\,\mathrm{id} \quad \text{in the sense of quadratic forms.}$$

Because of (14) $\Longrightarrow$ (15), we have for every $i\in\{1,\cdots,N\}$:

$$\forall\,(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N) \quad \mu(dx_i|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_N) \quad \text{satisfies LSI}(A_{ii}).$$

Finally, in view of (12), we have for every $i < j \in \{1, \cdots, N\}$:

$$\forall\, (x_1, \cdots, x_N) \quad |\nabla_i \nabla_j H(x_1, \cdots, x_N)| \;=\; -A_{ij}.$$

Remark 5 (Royer's Théorème 5.2.1) *In this remark, we compare our Theorem 1 to Royer's [R, Théorème 5.2.1]. In view of Lemma 6 below, Royer's hypothesis (H_2) on the coupling can be rephrased in our language (2) & (3) as*

$$\forall\, i \in \{1, \cdots, N\} \quad \sum_{j \neq i} \frac{\kappa_{ij}}{\rho_i} \;\leq\; \gamma, \tag{17}$$

for some constant $\gamma < 1$, where $\kappa_{ij} = \kappa_{ji}$. His hypothesis on the single–site conditional measures in our language (3) turns into

$$\forall\, i \in \{1, \cdots, N\} \quad \rho_i \;\geq\; \frac{2}{c_0}, \tag{18}$$

for some constant $c_0 > 0$. His result translates into

$$\mu(dx_1 \cdots dx_N) \quad \textit{satisfies LSI}\Big((1-\gamma)^2\, \frac{2}{c_0}\Big). \tag{19}$$

(In fact, there seems to be a typo in the statement of his result [R, (87)]. The inequality (19) is taken from [R, (92)] in Royer's proof, which is stronger than the actual statement [R, (87)] by a factor of two.)

On the other hand, (17) and (18) imply that the matrix A defined in (4) satisfies in the sense of quadratic forms

$$A \;\geq\; \frac{2}{c_0}(1-\gamma)\,\mathrm{id}. \tag{20}$$

Indeed, consider the smallest eigenvalue ρ of A. Let x denote a corresponding eigenvector. Let $i \in \{1, \cdots, N\}$ be such that $|x_i| = \max_j |x_j|$. W. l. o. g. we *may assume*

$$x_i \;=\; \max_j |x_j| \;=\; 1. \tag{21}$$

Then the i–th component of the identity $\rho x = Ax$ reads

$$\begin{aligned}
\rho &\overset{(21)}{=} \rho_i - \sum_{j\neq i} \kappa_{ij} x_j \\
&\geq \rho_i - \left(\sum_{j\neq i} \kappa_{ij}\right) \max_k |x_k| \\
&\overset{(21)}{=} \rho_i - \sum_{j\neq i} \kappa_{ij} \\
&\overset{(17)}{\geq} \rho_i - \gamma\, \rho_i \\
&\overset{(18)}{\geq} \frac{2}{c_0}(1-\gamma).
\end{aligned}$$

This establishes (20). Hence Theorem 1 implies that

$$\mu(dx_1 \cdots dx_N) \quad \textit{satisfies LSI}\left((1-\gamma)\,\frac{2}{c_0}\right),$$

which is stronger than Royer's by a factor of $(1-\gamma)$*.*

1.2 Two–scale theorem

We will apply Theorem 1 in conjunction with a "two–scale criterion" for LSI, implicitly contained in [GORV] Proposition 2 and [BB] Theorem 1.3. The two–scale criterion states that LSI for a conditional measure and the corresponding marginal may be combined – regardless of the interaction strength – to prove a global LSI. To be precise:

Theorem 2 *Let $\mu \in \mathcal{P}(X_1 \times X_2)$ be a probability measure with Hamiltonian*

H. Suppose that

$$\forall x_1 \in X_1,\ \mu(dx_2|x_1) \in \mathcal{P}(X_2) \ \textit{satisfies}\ LSI(\rho_2), \tag{22}$$

$$\bar{\mu}(dx_1) \in \mathcal{P}(X_1) \ \textit{satisfies}\ LSI(\bar{\rho}_1). \tag{23}$$

Then

$$\begin{aligned} &\mu \ \textit{satisfies}\ LSI(\rho) \ \textit{with} \\ \rho \ &\geq \ \frac{1}{2}\left(\rho_2 + (1+\alpha)\bar{\rho}_1 - \sqrt{(\rho_2+(1+\alpha)\bar{\rho}_1)^2 - 4\rho_2\bar{\rho}_1}\right), \end{aligned}$$

where

$$\alpha \ := \ \frac{1}{\rho_2}\frac{1}{\bar{\rho}_1}\sup_{X_1\times X_2} |\nabla_{X_1}\nabla_{X_2} H|^2. \tag{24}$$

Remark 6 (Product measures) *In the absence of coupling, i.e. $\alpha = 0$, we recover the product estimate (cf. Remark 1):*

$$\rho \ \geq \ \frac{1}{2}\left(\rho_2 + \bar{\rho}_1 - |\rho_2 - \bar{\rho}_1|\right) \ = \ \min\{\rho_2, \bar{\rho}_1\}.$$

Remark 7 (Linear algebra) *Theorem 1 formulates a sufficient condition for LSI in terms of linear algebra. Similarly, Theorem 2 has a connection with the linear algebra fact:*

$$\begin{aligned} \textit{Suppose}\ c \ &\geq \ \rho_2, \\ |b| \ &\leq \ h, \end{aligned}$$

and

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \ \geq \ \bar{\rho}_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \textit{in the sense of quadratic forms.}$$

Then

$$\begin{pmatrix} a & b \\ \\ b & c \end{pmatrix} \;\geq\; \rho\,\mathrm{id},$$

with

$$\rho \;=\; \frac{1}{2}\left(\rho_2 + (1+\alpha)\bar\rho_1 - \sqrt{(\rho_2 + (1+\alpha)\bar\rho_1)^2 - 4\rho_2\bar\rho_1}\right),$$

and

$$\alpha \;:=\; \frac{1}{\rho_2}\frac{1}{\bar\rho_1}h^2.$$

2 Two applications

2.1 Application I: Weak interactions

Our first example is a straightforward application of Theorem 1 in the case of non–convex potential and weak interactions. This problem is quite well–known in the literature and has been analyzed both by the methods of Zegarlinski, for instance in [R,BH], and by the method of Lu–Yau, for instance in [Yo]. A nice treatment is given in [L] Theorem 6.3. We include this example in order to illustrate that Theorem 1 makes the analysis particularly simple.

We consider nearest–neighbor interactions in two dimensions; without difficulty one can generalize to higher dimensions and either finite–range interactions or infinite–range interactions that decay sufficiently quickly.

Let X denote a two–dimensional periodic lattice and let the Gibbs measure

$\mu \in \mathcal{P}(X)$ have Hamiltonian

$$H(x) \; = \; \sum_i \psi(x_i) - \varepsilon \sum_{\langle i,j \rangle} x_i \, x_j, \tag{25}$$

where $\langle i, j \rangle$ represent nearest neighbor sites and the smooth potential ψ is a bounded perturbation of a Gaussian in the sense that

$$\psi \; = \; \frac{1}{2} x^2 + \delta\psi, \qquad \text{with} \qquad \sup_{\mathbb{R}} |\delta\psi(x)| < \infty. \tag{26}$$

Define $\Delta \; := \; \exp(-\mathrm{osc}_{\mathbb{R}} \, \delta\psi)$.

If the interaction is sufficiently weak in the sense that

$$\varepsilon \; \leq \; \frac{\Delta}{4}, \tag{27}$$

then μ satisfies LSI(ρ) with

$$\rho \; \geq \; \Delta - 4\varepsilon.$$

Let X_i denote a single site on the lattice, so that the full lattice is given by $X_1 \times \cdots \times X_N$. Then the single–site conditional measure

$$\mu_i(dx_i) \; := \; \mu(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N) \quad \text{satisfies LSI}(\Delta).$$

To see this, notice that the conditional measure is

$$\mu_i(dx_i) \; = \; \mathcal{Z}^{-1} \exp\Big(- \psi(x_i) + \varepsilon x_i \sum_{j;\langle j,i \rangle} x_j \Big) \, dx_i.$$

Here $j; \langle j, i \rangle$ is the set of sites j which are nearest neighbors of i. Thus, the measure is a perturbed Gaussian:

$$\mu_i(dx_i) \; = \; \mathcal{Z}^{-1} \exp\Big(- \frac{1}{2} x_i^2 + \varepsilon x_i \sum_{j;\langle j,i \rangle} x_j - \delta\psi(x_i) \Big) \, dx_i.$$

Since by the Bakry–Emery Principle the Gaussian satisfies LSI(1) (cf. [BE] and also Remark 2, below), we have by the Holley–Stroock principle [HS] that

$$\mu_i(dx_i) \quad \text{satisfies} \quad \mathrm{LSI}\Big(\exp(-\mathrm{osc}_{\mathbb{R}}\,\delta\psi)\Big). \tag{28}$$

Finally, observe that $|\nabla_i \nabla_j H| = \varepsilon$ for i and j nearest neighbors (there are four in two dimensions), and zero otherwise.

Thus, we may invoke Theorem 1. Given (27), the matrix $A \;\geq\; \rho\,\mathrm{id}$ with

$$\rho \;=\; \Delta - 4\varepsilon.$$

2.2 Application II: Order–one interactions with decay of correlations

Our second application considers order–one interactions under a decay of correlations condition. This "non–perturbative case" has been treated for instance in [Z3,SZ,Yo2]. We include this example in order to illustrate a more creative application of Theorem 1.

For simplicity, we consider product interactions and a two–dimensional lattice. Let X denote a two–dimensional N–periodic lattice generated by the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. We assume that the Gibbs measure $\mu \in \mathcal{P}(X)$ has Hamiltonian

$$H(x) \;=\; \sum_i \psi(x_i) + \sum_{i,j} J_{ij}\, x_i\, x_j, \tag{29}$$

where

$$J_{ii} \;\equiv\; 0, \qquad J_{ij} \;=\; J_{ji},$$

and the smooth potential ψ has the form

$$\psi \;=\; \frac{1}{4}x^4 + \delta\psi, \qquad \text{with} \qquad \sup_{\mathbb{R}} |\delta\psi''(x)| \le C. \tag{30}$$

(A common example is $x^4/4 - x^2/2$.)

We assume moreover that:

$$|J_{ij}| \;\le\; C \exp(-|i-j|/C) \tag{31}$$

and

$$|\langle x_i, x_j \rangle| \;\le\; C \exp(-|i-j|/C), \tag{32}$$

where $\langle \cdot, \cdot \rangle$ denotes the expectation with respect to the conditional measure obtained by conditioning on any other sites of the lattice.

Then there exists a positive constant ρ independent of N such that

$$\mu \qquad \text{satisfies} \qquad \mathrm{LSI}(\rho).$$

We begin by noting the natural consequences of our assumptions. First:

Lemma 1 *For every $C < \infty$ there exists $C_b < \infty$ such that if*

$$\psi \;=\; \frac{1}{4}x^4 + \delta\psi, \qquad \textit{with} \qquad \sup_{\mathbb{R}} |\delta\psi''(x)| \;\le\; C, \tag{33}$$

then ψ can be written

$$\psi(x) = \psi_c(x) + \psi_b(x),$$

where the convex function ψ_c and the bounded function ψ_b satisfy

$$\psi_c''(x) \ge 1, \qquad \sup_{\mathbb{R}} |\psi_b(x)| \le C_b. \tag{34}$$

Together with our assumption (31), this implies:

Lemma 2 (Single–site conditionals) *There exists* $\rho_s > 0$ *such that for any subset* $Y \subset X$*, the marginal* $\bar{\mu} \in \mathcal{P}(X \setminus Y)$ *is such that for any* $i \in X \setminus Y$*, the single–site conditional measure*

$$\bar{\mu}(dx_i \mid j \notin Y, j \neq i) \qquad \text{satisfies} \qquad \text{LSI}(\rho_s).$$

Furthermore, the coarse–grained Hamiltonian $\bar{H}$ of $\bar{\mu}$ inherits exponential decay from (31) and (32):

Lemma 3 (Coarse–grained Hamiltonian) *Given* (31) *and* (32)*, there exists a constant* C_H *such that the following holds: For any* $Y, Z \subset X$*, the Hamiltonian* $\bar{H}$ *corresponding to* $\bar{\mu}(\Pi_i dx_i, i \in Z \mid j \notin (Y \cup Z))$ *satisfies*

$$|\nabla_{i_1} \nabla_{i_2} \bar{H}| \;\leq\; C_H \exp(-|i_1 - i_2|/C_H) \qquad \forall\, i_1, i_2 \in Z. \tag{35}$$

In particular, consider the K–sublattice (cf. Figure 1), defined as

$$X_K := \{n\, K\, \mathbf{e}_1 + m\, K\, \mathbf{e}_2,\; n, m \in \mathbb{Z}\} \cap X.$$

Theorem 1 together with Lemmas 2 and 3 imply:

Lemma 4 (K–sublattice) *Given any* $\varepsilon > 0$*, there exists* K *sufficiently large so that: For any subset* $Y \subset X \setminus X_K$ *and the measure* $\bar{\mu}^Y \in \mathcal{P}(X_K)$ *which is averaged over* Y *and conditioned on* $X \setminus (Y \cup X_K)$*,*

$$\bar{\mu}^Y \qquad \text{satisfies} \qquad \text{LSI}(\rho_s - \varepsilon). \tag{36}$$

To prove LSI for the full measure μ, we now apply the two scale criterion from Theorem 2 to a sequence of conditional and marginal measures (cf. Figures 1 through 4).

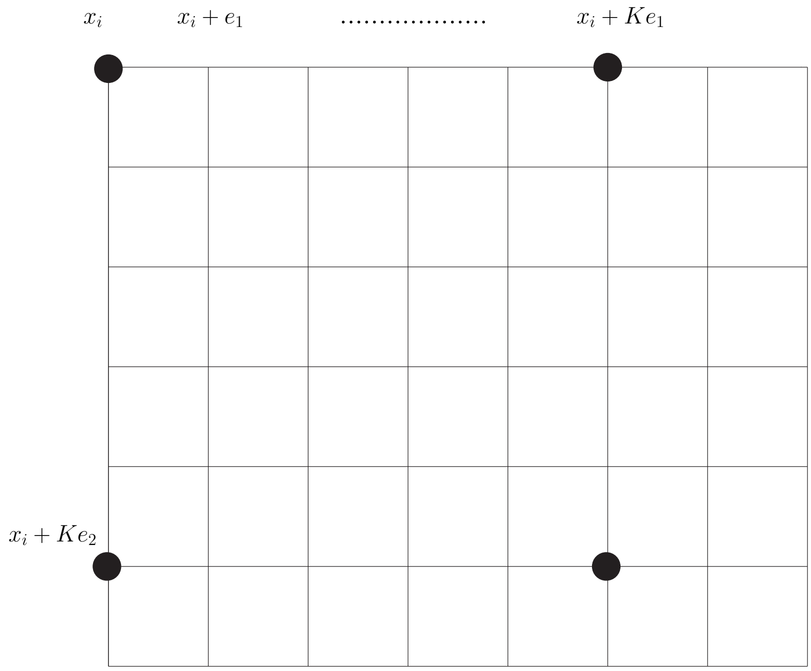

Fig. 1. We establish LSI for $\mu_c^{(1)}$, the measure that is "active" on every K^{th} site (black circles) and conditioned on the spin at every other site (no circles drawn).

Before explaining the sequence, it is convenient to introduce some notation. Let X_K^ℓ for $\ell = 0, \ldots, (K^2 - 1)$ be an enumeration of the translates of X_K:

$$X_K + c_1\, \mathbf{e}_1 + c_2 \mathbf{e}_2, \qquad \text{for } c_1, c_2 \in 1, \ldots, K-1,$$

and let

$$Y_K^n \;=\; \cup_{\ell=0}^n X_K^\ell.$$

By definition, $X_K^0 = Y_K^0 = X_K$.

Finally, define

$$\mu_c^{(n)} \;:=\; \mu\left(\Pi_i dx_i, i \in Y_K^{n-1} \mid j \notin Y_K^{n-1}\right)$$

and

$$\bar{\mu}^{(n)} \;:=\; \bar{\mu}\left(\Pi_i dx_i, i \in X_K^n \mid x_j, j \notin Y_K^n\right).$$

By Lemma 4 with $Y = Y_K^0 = X_K$, $\mu_c^{(1)}$ satisfies LSI$(\rho_s - \varepsilon)$. By Lemma 4 with $Y = Y_K^1$, $\bar{\mu}^{(1)}$ satisfies LSI$(\rho_s - \varepsilon)$. By Theorem 2, these two ingredients imply

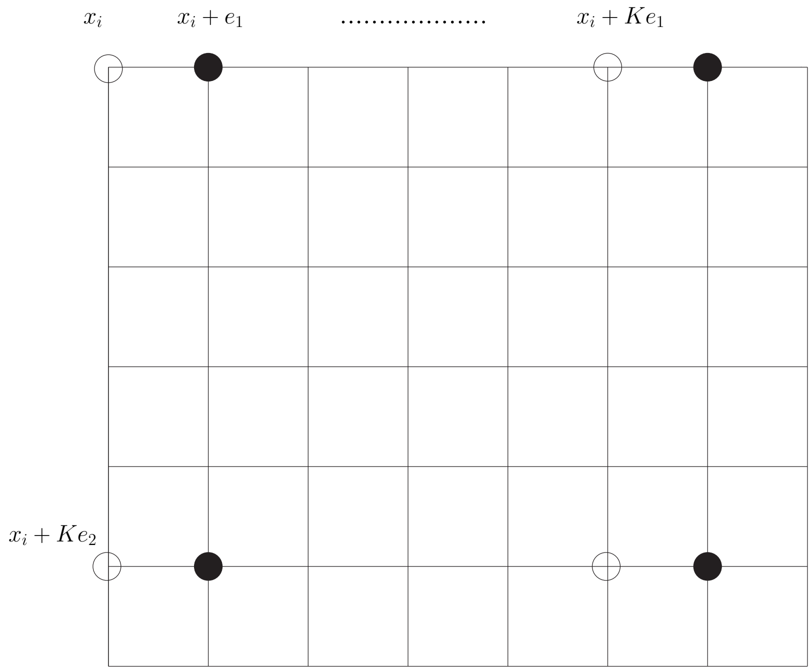

Fig. 2. We establish LSI for $\bar{\mu}^{(1)}$, the measure that is "active" on the black circles, averaged over the open circles, and conditioned on the spin at the other sites.

LSI for $\mu_c^{(2)}$. (See Figures 1, 2, and 3 for an illustration.)

We now use this information, Lemma 4 for $\bar{\mu}^{(2)}$, and Theorem 2 to conclude LSI for $\mu_c^{(3)}$. We continue in this way until having proved LSI for

$$\mu_c^{(K^2)} = \mu.$$

2.2.1 Proofs of the lemmas

Proof of Lemma 1. Below, C is an order–one constant which may change from line to line but depends at most on $\sup_{\mathbb{R}} |\delta\psi''(x)|$.

First we remark that by (33), we may write

$$\delta\psi(x) \;=\; \delta\psi(0) + \delta\psi'(0)x + \xi(x),$$

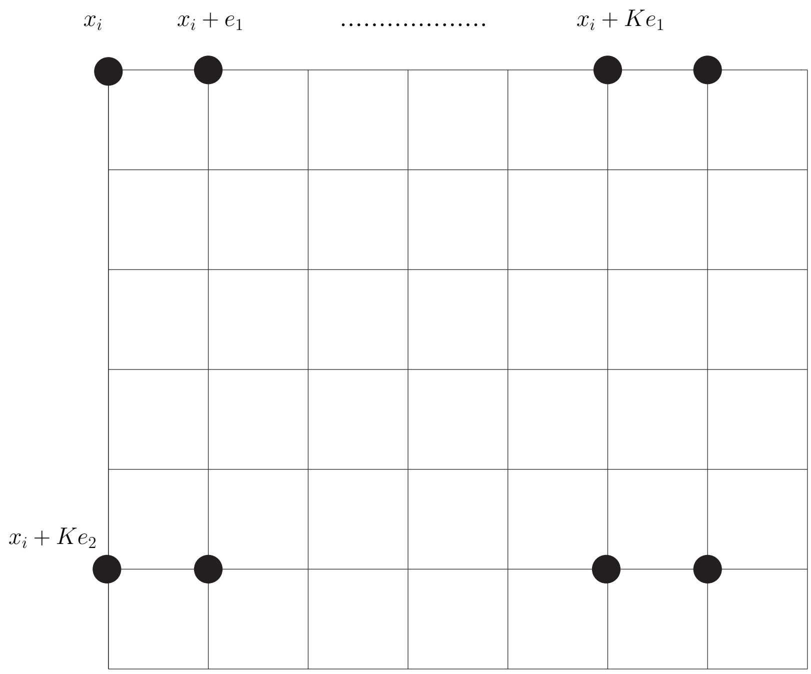

Fig. 3. We use the two–scale criterion from Theorem 2 to deduce from LSI for $\mu_c^{(1)}$ and $\bar{\mu}^{(1)}$ the LSI for $\mu_c^{(2)}$, the measure that is “active” on the black circles.

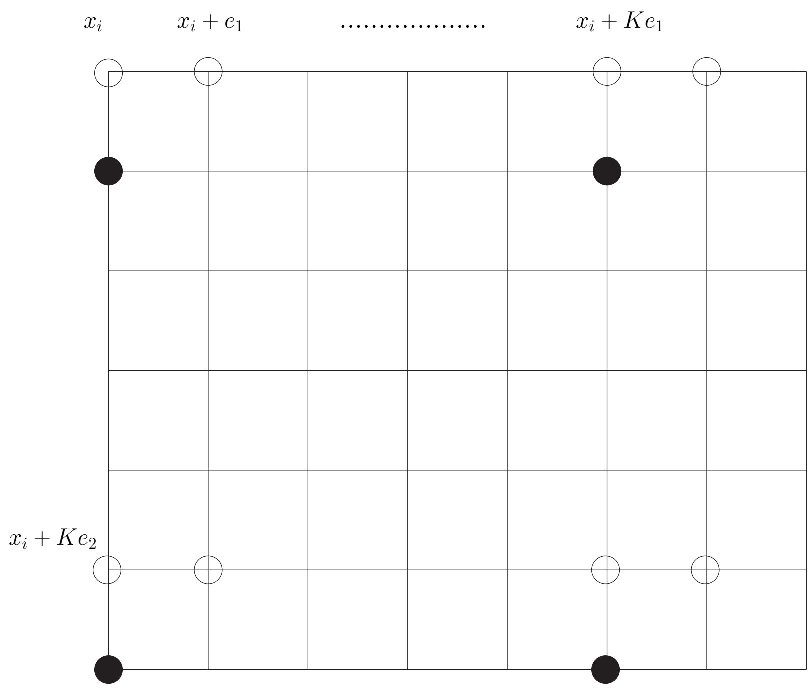

Fig. 4. This illustration depicts $\bar{\mu}^{(2)}$.

where ξ is a function such that

$$\xi(0) \;=\; 0, \qquad \xi'(0) \;=\; 0, \qquad \sup_{\mathbb{R}} |\xi''(x)| \;\leq\; C. \tag{37}$$

Let

$$\begin{aligned} \psi_c(x) \;&=\; \psi(x) + \frac{x^2}{2}\,\phi_R(x) - \xi(x)\,\phi_R(x), \\ \psi_b(x) \;&=\; -\frac{x^2}{2}\,\phi_R(x) + \xi(x)\,\phi_R(x), \end{aligned}$$

where $\phi_R(x) = \phi(x/R)$ and ϕ is a cut–off function such that

$$\phi(x) \equiv 1 \;\text{ for }\; x \in [-1,1], \quad \phi(x) \equiv 0 \;\text{ for }\; x \notin (-2,2), \quad |\phi(x)| \;\leq\; 1.$$

We will show that by choosing R sufficiently large, (34) is satisfied. Notice that

$$|\phi_R(x)| \;\leq\; 1, \qquad |\phi_R'(x)| \;\leq\; \frac{C}{R}, \qquad |\phi_R''(x)| \;\leq\; \frac{C}{R^2}. \tag{38}$$

We begin with ψ_c, for which:

$$\begin{aligned} \psi_c''(x) \;=\; & 3\,x^2 + \xi''(x) + \phi_R(x) + 2\,x\,\phi_R'(x) + \frac{x^2}{2}\,\phi_R''(x) \\ & - \xi(x)\,\phi_R''(x) - 2\,\xi'(x)\,\phi_R'(x) - \xi''(x)\,\phi_R(x). \end{aligned} \tag{39}$$

We have that:

$$\begin{aligned} |x| \;\leq\; R \quad &\Rightarrow \quad \psi_c''(x) \;=\; 3\,x^2 + \xi''(x) + 1 - \xi''(x) \;\geq\; 1, \\ |x| \;\geq\; 2R, \quad &\Rightarrow \quad \psi_c''(x) \;\geq\; 3(2R)^2 - \sup_{\mathbb{R}} |\xi''(x)| \;\overset{(33)}{\geq}\; 1, \end{aligned}$$

for $R^2 \geq (1+C)/12$. Now consider $R \;<\; |x| \;<\; 2R$. Notice that by (37),

$$\sup_{[-2R,2R]} |\xi'(x)| \;\leq\; C\,R, \qquad \sup_{[-2R,2R]} |\xi(x)| \;\leq\; C\,R^2. \tag{40}$$

Using (37), (38), and (40) in (39) gives:

$$
\begin{aligned}
\psi_c''(x) &\geq 3R^2 - C - 2(2R)\frac{C}{R} - \frac{(2R)^2}{2}\frac{C}{R^2} - CR^2\frac{C}{R^2} - 2CR\frac{C}{R} - C \\
&\geq 3R^2 - C \ \geq \ 1,
\end{aligned}
$$

for $R^2 \geq (1+C)/3$.

Finally, notice

$$
|\psi_b(x)| \ \leq \ \frac{(2R)^2}{2} + \sup_{[-2R,2R]} |\xi(x)| \ \overset{(40)}{\leq} \ CR^2 \ =: \ C_b.
$$

Proof of Lemma 2.

The Hamiltonian corresponding to $\bar{\mu}(dx_i \mid j \notin Y, j \neq i)$ is of the form

$$
\bar{H}(x_i) \ = \ \psi(x_i) + \text{linear} + \bar{h}(x_i),
$$

where

$$
\bar{h}(x_i) \ := \ -\log \int \exp\left(-\sum_{\ell \in Y} \psi(x_\ell) - \sum_{\ell \in Y} \sum_{k \in X} J_{\ell k}\, x_\ell\, x_k \right) \prod_{\ell \in Y} dx_\ell. \tag{41}
$$

We claim:

$$
|\bar{h}''(x_i)| \ \leq \ C. \tag{42}
$$

By Lemma 1, (30) and (42) imply that we may write

$$
\bar{H}(x_i) \ = \ \psi_c(x_i) + \psi_b(x_i)
$$

with

$$
\psi_c''(x) \geq 1, \qquad \sup_{\mathbb{R}} |\psi_b(x)| \leq C_b.
$$

By the Bakry–Emery principle,

$$\mathcal{Z}^{-1}\exp\left(-\psi_c(x)+\text{linear}\right) \quad \text{satisfies} \quad \text{LSI}(1).$$

Together with the Holley–Stroock principle, this implies that

$$\bar{\mu}(dx_i \mid x_j,\, j\notin Y, j\neq i) \quad \text{satisfies} \quad \text{LSI}(\exp(-\text{osc}_{\mathbb{R}}\psi_b)).$$

Thus, it remains only to establish (42). For this we calculate:

$$\bar{h}'(x_i) \;=\; \left\langle \sum_{\ell\in Y} J_{\ell\, i} x_\ell \right\rangle, \qquad \bar{h}''(x_i) \;=\; -\text{var}\left(\sum_{\ell\in Y} J_{\ell\, i} x_\ell\right), \tag{43}$$

where $\langle\cdot\rangle$ and var denote the expectation and variance with respect to

$$\mu(\Pi_{\ell\in Y} dx_\ell \mid x_j, j\notin Y). \tag{44}$$

We remark:

$$\begin{aligned}\text{var}\left(\sum_{\ell\in Y} J_{\ell\, i}\, x_\ell\right) \;&=\; \sum_{\ell_1,\ell_2\in Y} J_{\ell_1\, i} J_{\ell_2\, i}\,\text{cov}(x_{\ell_1}, x_{\ell_2})\\ &\overset{(32)}{\leq}\; C\sum_{\ell_1\in Y}|J_{\ell_1\, i}|\sum_{\ell_2\in Y}|J_{\ell_2\, i}|.\end{aligned}$$

Each sum on the right–hand side is bounded by:

$$\begin{aligned}\sum_{\ell\in Y}|J_{\ell\, i}| \;&\overset{(31)}{\leq}\; C\sum_{\ell\in Y}\exp(-|\ell-i|/C)\\ &\leq\; C\sum_{\ell\in\mathbb{Z}^2}\exp(-|\ell|/C)\\ &\leq\; C\sum_{\ell\in\mathbb{Z}^2}\exp(-|\ell|_\infty/C)\\ &\leq\; C\sum_{n=1}^{\infty} 8\, n\,\exp(-\,n/C) \;\leq\; C.\end{aligned} \tag{45}$$

Thus,

$$\operatorname{var}\left(\sum_{\ell\in Y} J_{\ell\, i}\, x_\ell\right) \;\le\; C,$$

which, together with (43), establishes (42).

Proof of Lemma 3.

By direct calculation and the triangle inequality, it follows that:

$$\begin{aligned}
\left|\nabla_{i_1}\nabla_{i_2}\bar{H}\right| &= \left|J_{i_1\, i_2} - \operatorname{var}\left(\sum_{\ell\in Y} J_{i_1\,\ell}\, x_\ell \sum_{\ell\in Y} J_{i_2\,\ell}\, x_\ell\right)\right| \\
&= \left|J_{i_1\, i_2} - \sum_{\ell_1,\ell_2\in Y} J_{i_1\,\ell_1}\operatorname{cov}(x_{\ell_1},\, x_{\ell_2}) J_{\ell_2\, i_2}\right| \\
&\le |J_{i_1\, i_2}| + \left|\sum_{\ell_1,\ell_2\in Y} J_{i_1\,\ell_1}\operatorname{cov}(x_{\ell_1},\, x_{\ell_2}) J_{\ell_2\, i_2}\right|, \qquad (46)
\end{aligned}$$

where var and cov denote the variance and covariance with respect to the conditional measure (44). We claim that there exists $C_2 \ge C$ such that the second term satisfies

$$\left|\sum_{\ell_1,\ell_2\in Y} J_{i_1\,\ell_1}\operatorname{cov}(x_{\ell_1},\, x_{\ell_2}) J_{\ell_2\, i_2}\right| \;\le\; C_2 \exp(-|i_1 - i_2|/C_2). \qquad (47)$$

The combination of (46), (31), and (47) implies (35) with $C_H := 2C_2$.

Thus, we need only show (47). According to our assumptions,

$$\begin{aligned}
&\left|\sum_{\ell_1,\ell_2\in Y} J_{i_1\,\ell_1}\operatorname{cov}(x_{\ell_1},\, x_{\ell_2}) J_{\ell_2\, i_2}\right| \\
&\le C \sum_{\ell_1,\ell_2\in Y} \exp(-(|i_1-\ell_1| + |\ell_1-\ell_2| + |\ell_2 - i_2|)/C). \qquad (48)
\end{aligned}$$

By the triangle inequality,

$$|i_1 - i_2| \;\le\; |i_1-\ell_1| + |\ell_1-\ell_2| + |\ell_2 - i_2|,$$

we deduce

$$\begin{aligned}\frac{1}{2}\left(|i_1 - i_2| + |i_1 - \ell_1| + |i_2 - \ell_2|\right) &\leq |i_1 - \ell_1| + \frac{1}{2}|\ell_1 - \ell_2| + |i_2 - \ell_2| \\ &\leq |i_1 - \ell_1| + |\ell_1 - \ell_2| + |i_2 - \ell_2|.\end{aligned}$$

Hence,

$$\begin{aligned}&\exp(-(|i_1 - \ell_1| + |\ell_1 - \ell_2| + |\ell_2 - i_2|)/C) \\ &\leq \exp(-(|i_1 - i_2| + |i_1 - \ell_1| + |i_2 - \ell_2|)/(2C)).\end{aligned}$$

Hence,

$$\begin{aligned}&\sum_{\ell_1, \ell_2 \in Y} \exp(-(|i_1 - \ell_1| + |\ell_1 - \ell_2| + |\ell_2 - i_2|)/C) \\ &\leq \exp(-|i_1 - i_2|/(2C)) \sum_{\ell_1 \in Y} \exp(-|\ell_1 - i_1|/(2C)) \sum_{\ell_2 \in Y} \exp(-|i_2 - \ell_2|/(2C)).\end{aligned}$$

Bounding the sums as in (45) and substituting into (48) returns (47).

PROOF OF LEMMA 4.

According to Lemma 2, for any i in X_K, the single–site conditional measure

$$\bar{\mu}^Y(dx_i \mid j \in X_K, j \neq i) \qquad \text{satisfies} \qquad \text{LSI}(\rho_s). \tag{49}$$

According to Lemma 3, the sum of interactions for any site i is bounded by:

$$\begin{aligned}
\sum_{\substack{k\in X_K,\\ k\neq i}} |\nabla_i\nabla_k H| &\leq C \sum_{\substack{k\in X_K,\\ k\neq i}} \exp(-|i-k|/C) \\
&\leq C \sum_{k\in\mathbb{Z}^2\setminus\{0\}} \exp(-K|k|/C) \\
&\leq C \sum_{k\in\mathbb{Z}^2\setminus\{0\}} \exp(-K|k|_\infty/C) \\
&= C \sum_{\ell=1}^{\infty} 8\,\ell\,\exp(-K\ell/C) \\
&< \varepsilon \qquad (50)
\end{aligned}$$

for K sufficiently large.

By Theorem 1, (49) and (50) combine to give (36).

3 Auxiliary results for Theorem 1

At the core of Theorem 1 is a covariance estimate stated in Lemma 5 below. It goes back to Bodineau & Helffer. Ledoux gave a very efficient proof in [L, Proposition 2.2] . We give yet a different proof which mimics the proof of Talagrand's inequality given in [OV].

Lemma 5 *Let $\mu(dx)$ be a probability measure on the Euclidean space X. We assume that there exists $\rho > 0$ such that*

$$\mu \quad \text{satisfies LSI}(\rho). \qquad (51)$$

Then we have for arbitrary $f(x) \geq 0$ and $g(x)$:

$$
\begin{aligned}
&\left|\int g\,f\,d\mu - \int g\,d\mu \int f\,d\mu\right| \\
&\leq \sup_x |\nabla g| \left(\frac{2}{\rho}\int f\,d\mu \left(\int \Phi(f)\,d\mu - \Phi\left(\int f\,d\mu\right)\right)\right)^{1/2} \\
&\leq \sup_x |\nabla g|\,\frac{1}{\rho}\left(\int f\,d\mu \int \frac{1}{f}|\nabla f|^2\,d\mu\right)^{1/2}. \qquad (52)
\end{aligned}
$$

We also need a linearized version of Lemma 5:

Corollary 1 *Let $\mu(dx)$ be a probability measure on the Euclidean space X. We assume that there exists $\rho > 0$ such that*

$$\mu \quad \textit{satisfies LSI}(\rho).$$

Then we have for arbitrary $g(x)$ and $h(x)$:

$$\left|\int g\,h\,d\mu - \int g\,d\mu \int h\,d\mu\right| \;\leq\; \frac{1}{\rho}\,\sup_x |\nabla g|\,\sup_x |\nabla h|. \qquad (53)$$

Lemma 5 will be used to establish the following version of [R, Hypothesis H2, p.91].

Lemma 6 *Let X_1, X_2 be two Euclidean spaces and $\mu(dx_1 dx_2)$ a probability measure on the product space $X_1 \times X_2$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

We assume that there exists $\kappa_{12} < \infty$ such that the Hamiltonian $H(x_1, x_2) = -\ln \frac{d\mu}{d\mathcal{L}}$ satisfies

$$\forall\;(x_1, x_2) \quad |\nabla_1 \nabla_2 H(x_1, x_2)| \;\leq\; \kappa_{12}.$$

We assume that there exists $\rho_2 > 0$ such that we have for the conditional measure

$$\forall x_1 \quad \mu(dx_2|x_1) \quad \textit{satisfies LSI}(\rho_2).$$

For arbitrary $f(x_1,x_2)\ge 0$*, consider*

$$\bar f(x_1) \;=\; \int f(x_1,x_2)\,\mu(dx_2|x_1). \tag{54}$$

Then we obtain for the marginal $\bar\mu(dx_1)$

$$\begin{aligned}&\Big(\int \frac{1}{2\bar f}|\nabla_1\bar f|^2\,\bar\mu(dx_1)\Big)^{1/2}\\ &\le\;\Big(\int \frac{1}{2f}|\nabla_1 f|^2\,d\mu\Big)^{1/2}\;+\;\frac{\kappa_{12}}{\rho_2}\,\Big(\int \frac{1}{2f}|\nabla_2 f|^2\,d\mu\Big)^{1/2}.\end{aligned}$$

The proof of the following lemma, which is based on Lemma 6, amounts to the Lu–Yau martingale method [LY] in the case of only two sites.

Lemma 7 *Let* X_1,X_2 *be two Euclidean spaces and* $\mu(dx_1dx_2)$ *a probability measure on the product space* $X_1\times X_2$ *with a smooth positive Lebesgue density* $\frac{d\mu}{d\mathcal{L}}$.

We assume that there exists $\kappa_{12}<\infty$ *such that the Hamiltonian* $H(x_1,x_2)=-\ln\frac{d\mu}{d\mathcal{L}}$ *satisfies*

$$\forall\;(x_1,x_2)\quad |\nabla_1\nabla_2 H(x_1,x_2)|\;\le\;\kappa_{12}.$$

We assume that there exist $\rho_2,\bar\rho_1>0$ *such that we have for the conditional measure and the marginal*

$$\forall x_1\quad \mu(dx_2|x_1)\ \text{satisfies LSI}(\rho_2), \tag{55}$$

$$\bar\mu(dx_1)\ \text{satisfies LSI}(\bar\rho_1). \tag{56}$$

Then we obtain for the marginal $\bar\mu(dx_2)$

$$\bar\mu(dx_2)\quad \text{satisfies LSI}(\bar\rho_2)$$

with

$$\frac{1}{\bar{\rho}_2} \;\leq\; \frac{1}{\rho_2} + \frac{1}{\bar{\rho}_1}\frac{\kappa_{12}^2}{\rho_2^2}.$$

The following statement is a simple consequence of Lemma 7. Alternatively, it can be obtained by Zegarlinski's iterative argument, which is outlined for instance in [GZ, Section 5.2].

Corollary 2 *Let X_1, X_2 be two Euclidean spaces and $\mu(dx_1 dx_2)$ a probability measure on the product space $X_1 \times X_2$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

We assume that there exists $\kappa_{12} < \infty$ such that the Hamiltonian $H(x_1, x_2) = -\ln \frac{d\mu}{d\mathcal{L}}$ satisfies

$$\forall\ (x_1, x_2) \quad |\nabla_1 \nabla_2 H(x_1, x_2)| \;\leq\; \kappa_{12}.$$

We assume that there exist $\rho_1, \rho_2 > 0$ such that we have for the conditional measures

$$\begin{array}{ll} \forall x_2 & \mu(dx_1|x_2) \text{ satisfies } LSI(\rho_1), \\ \forall x_1 & \mu(dx_2|x_1) \text{ satisfies } LSI(\rho_2). \end{array}$$

We assume that

$$\rho_1\, \rho_2 - \kappa_{12}^2 \;>\; 0. \tag{57}$$

Then we obtain for the marginal $\bar{\mu}(dx_1)$

$$\bar{\mu}(dx_1) \quad \text{satisfies } LSI(\bar{\rho}_1)$$

with

$$\bar{\rho}_1 \;\geq\; \rho_1 - \frac{\kappa_{12}^2}{\rho_2}.$$

We will also need the following consequence of Corollary 1.

Lemma 8 *Let X_1, X_2, X_3 be Euclidean spaces and $\mu(dx_1 dx_2 dx_3)$ a probability measure on the product space $X_1 \times X_2 \times X_3$ with a smooth positive Lebesgue density $\frac{d\mu}{d\mathcal{L}}$.*

We assume that for $i < j \in \{1, 2, 3\}$ there exists $\kappa_{ij} < \infty$ such that the Hamiltonian $H(x_1, x_2, x_3) = -\ln \frac{d\mu}{d\mathcal{L}}$ satisfies

$$\forall \, (x_1, x_2, x_3) \quad |\nabla_i \nabla_j H(x_1, x_2, x_3)| \ \leq \ \kappa_{ij}.$$

We assume that there exists $\rho_3 > 0$ such that we have for the conditional measures

$$\forall (x_1, x_2) \quad \mu(dx_3 | x_1, x_2) \quad \text{satisfies } LSI(\rho_3).$$

Consider the Hamiltonian $\bar{H}(x_1, x_2)$ belonging to the marginal $\bar{\mu}(dx_1, dx_2)$, i.e.

$$\bar{H}(x_1, x_2) \ = \ -\ln \int \exp(-H(x_1, x_2, x_3)) \, dx_3.$$

It satisfies

$$\forall \, (x_1, x_2) \quad |\nabla_1 \nabla_2 \bar{H}(x_1, x_2)| \ \leq \ \bar{\kappa}_{12}.$$

with

$$\bar{\kappa}_{12} \ \leq \ \kappa_{12} + \frac{\kappa_{13} \, \kappa_{23}}{\rho_3}. \tag{58}$$

Finally, we need an elementary result from linear algebra, which we reproduce for convenience.

Lemma 9 *Consider a symmetric and positive definite matrix A with*

$$A_{ij} \ \leq \ 0 \quad for\, i < j \in \{1, \cdots, N\}.$$

Then the inverse matrix A^{-1} satisfies

$$(A^{-1})_{ij} \;\geq\; 0 \quad \mathit{for}\, i,j \in \{1,\cdots,N\}.$$

4 Proofs

Proof of Lemma 5.

Without loss of generality, we may assume

$$\int f\, d\mu \;=\; 1.$$

The second inequality in (52) follows from the first and (51). In order to prove the first, we introduce the semigroup P_t related to μ and defined by

$$P_0 f = f, \tag{59}$$

$$\forall\; g(x) \quad \frac{d}{dt}\int g\, P_t f\, d\mu = -\int \nabla g \cdot \nabla P_t f\, d\mu. \tag{60}$$

All we need to know is

$$\int P_t f\, d\mu = \int f\, d\mu \;=\; 1, \tag{61}$$

$$\frac{d}{dt}\int \Phi(P_t f)\, d\mu = -\int \frac{1}{P_t f}|\nabla P_t f|^2\, d\mu, \tag{62}$$

$$P_\infty f \;:=\; \lim_{t\uparrow\infty} P_t f = \int f\, d\mu \;=\; 1. \tag{63}$$

Indeed, the left–hand side of (52) can be reformulated as

$$\begin{aligned}
\int g\, f\, d\mu - \int g\, d\mu \int f\, d\mu &\overset{(59),(63)}{=} \int g\,(P_0 f - P_\infty f)\, d\mu \\
&= \int_0^\infty \frac{d}{dt}\int g\, P_t f\, d\mu\, dt \\
&\overset{(60)}{=} -\int_0^\infty \int \nabla g \cdot \nabla P_t f\, d\mu\, dt.
\end{aligned}$$

This yields the estimate

$$
\begin{aligned}
&\left|\int g\, f\, d\mu - \int g\, d\mu \int f\, d\mu\right| \\
&\leq \sup_x |\nabla g| \int_0^\infty \int |\nabla P_t f|\, d\mu\, dt \\
&\leq \sup_x |\nabla g| \int_0^\infty \left(\int P_t f\, d\mu \int \frac{1}{P_t f} |\nabla P_t f|^2\, d\mu\right)^{1/2} dt \\
&\overset{(61)}{=} \sup_x |\nabla g| \int_0^\infty \left(\int \frac{1}{P_t f} |\nabla P_t f|^2\, d\mu\right)^{1/2} dt.
\end{aligned}
$$

It remains to estimate the last term:

$$
\begin{aligned}
&\int_0^\infty \left(\int \frac{1}{P_t f}|\nabla P_t f|^2\, d\mu\right)^{1/2} dt \\
&= \int_0^\infty \left(\int \frac{1}{P_t f}|\nabla P_t f|^2\, d\mu\right)^{-1/2} \int \frac{1}{P_t f}|\nabla P_t f|^2\, d\mu\, dt \\
&\overset{(51),(61)}{\leq} \left(\frac{1}{2\rho}\right)^{1/2} \int_0^\infty \left(\int \Phi(P_t f)\, d\mu\right)^{-1/2} \int \frac{1}{P_t f}|\nabla P_t f|^2\, d\mu\, dt \\
&\overset{(62)}{=} -\left(\frac{1}{2\rho}\right)^{1/2} \int_0^\infty \left(\int \Phi(P_t f)\, d\mu\right)^{-1/2} \frac{d}{dt}\int \Phi(P_t f)\, d\mu\, dt \\
&= -\left(\frac{2}{\rho}\right)^{1/2} \int_0^\infty \frac{d}{dt}\left(\int \Phi(P_t f)\, d\mu\right)^{1/2} dt \\
&= \left(\frac{2}{\rho}\right)^{1/2} \left(\left(\int \Phi(P_0 f)\, d\mu\right)^{1/2} - \left(\int \Phi(P_\infty f)\, d\mu\right)^{1/2}\right) \\
&\overset{(59),(63)}{=} \left(\frac{2}{\rho}\right)^{1/2} \left(\int \Phi(f)\, d\mu\right)^{1/2}.
\end{aligned}
$$

Proof of Corollary 1.

Let $g(x)$ and $h(x)$ be given. We may assume that h is bounded so that for sufficiently small $\epsilon > 0$ we have

$$f(x) \;:=\; 1 + \epsilon\, h(x) \;\geq\; 0.$$

We then may apply Lemma 5 to f and g, which yields

$$\begin{aligned}
&\epsilon \left| \int g\,h\,d\mu - \int g\,d\mu \int h\,d\mu \right| \\
&\le \frac{1}{\rho} \sup_x |\nabla g| \left(1 + \epsilon \int h\,d\mu\right)^{1/2} \left(\epsilon^2 \int \frac{1}{1+\epsilon\,h} |\nabla h|^2\,d\mu\right)^{1/2}.
\end{aligned}$$

Dividing by ϵ and letting it tend to zero yields

$$\left| \int g\,h\,d\mu - \int g\,d\mu \int h\,d\mu \right| \;\le\; \frac{1}{\rho} \sup_x |\nabla g| \left(\int |\nabla h|^2\,d\mu \right)^{1/2},$$

which is a stronger version of (53).

Proof of Lemma 6.

From the representation

$$\bar f(x_1) \;=\; \int f(x_1,x_2)\,\mu(dx_2|x_1) \;=\; \frac{\int f(x_1,x_2)\,\exp(-H(x_1,x_2))\,dx_2}{\int \exp(-H(x_1,x_2))\,dx_2},$$

we deduce the formula

$$\begin{aligned}
\nabla_1 \bar f(x_1) = &\frac{\int \nabla_1 f(x_1,x_2)\,\exp(-H(x_1,x_2))\,dx_2}{\int \exp(-H(x_1,x_2))\,dx_2} \\
&- \frac{\int f(x_1,x_2)\,\nabla_1 H(x_1,x_2)\,\exp(-H(x_1,x_2))\,dx_2}{\int \exp(-H(x_1,x_2))\,dx_2} \\
&+ \frac{\int f(x_1,x_2)\,\exp(-H(x_1,x_2))\,dx_2}{\int \exp(-H(x_1,x_2))\,dx_2} \\
&\times \frac{\int \nabla_1 H(x_1,x_2)\,\exp(-H(x_1,x_2))\,dx_2}{\int \exp(-H(x_1,x_2))\,dx_2} \\
= &\int \nabla_1 f(x_1,x_2)\,\mu(dx_2|x_1) \\
&- \Big(\int f(x_1,x_2)\,\nabla_1 H(x_1,x_2)\,\mu(dx_2|x_1) \\
&\quad - \int f(x_1,x_2)\,\mu(dx_2|x_1) \int \nabla_1 H(x_1,x_2)\,\mu(dx_2|x_1) \Big).
\end{aligned}$$

Hence Lemma 5, applied to $\mu(dx_2|x_1)$, $f(x_1,x_2)$, and $g(x_2) = \nabla_1 H(x_1,x_2)$ for fixed x_1, yields

$$
\begin{aligned}
|\nabla_1 \bar{f}(x_1)| \leq & \int |\nabla_1 f(x_1,x_2)| \, \mu(dx_2|x_1) \\
& + \frac{1}{\rho_2} \sup_{x_2} |\nabla_2 \nabla_1 H(x_1,x_2)| \, \left(\int f(x_1,x_2) \, \mu(dx_2|x_1) \right)^{1/2} \\
& \quad \left(\int \frac{1}{f(x_1,x_2)} |\nabla_2 f(x_1,x_2)|^2 \, \mu(dx_2|x_1) \right)^{1/2} \\
\leq & \, (\bar{f}(x_1))^{1/2} \left(\left(\int \frac{1}{f(x_1,x_2)} |\nabla_1 f(x_1,x_2)|^2 \, \mu(dx_2|x_1) \right)^{1/2} \right. \\
& \left. + \frac{\kappa_{12}}{\rho_2} \left(\int \frac{1}{f(x_1,x_2)} |\nabla_2 f(x_1,x_2)|^2 \, \mu(dx_2|x_1) \right)^{1/2} \right).
\end{aligned}
$$

We rewrite this inequality as

$$
\begin{aligned}
& \left(\frac{1}{\bar{f}(x_1)} |\nabla_1 \bar{f}(x_1)|^2 \right)^{1/2} \\
\leq & \left(\int \frac{1}{f(x_1,x_2)} |\nabla_1 f(x_1,x_2)|^2 \, \mu(dx_2|x_1) \right)^{1/2} \\
& + \frac{\kappa_{12}}{\rho_2} \left(\int \frac{1}{f(x_1,x_2)} |\nabla_2 f(x_1,x_2)|^2 \, \mu(dx_2|x_1) \right)^{1/2}.
\end{aligned}
$$

The triangle inequality in $L^2(\bar{\mu}(dx_1))$ yields the desired result.

Proof of Lemma 7.

Let an arbitrary $f(x_2) \geq 0$ be given. We set for abbreviation

$$
\bar{f}(x_1) \; = \; \int f(x_2) \, \mu(dx_2|x_1). \tag{64}
$$

We split the left–hand side of the LSI as follows:

$$
\begin{aligned}
&\int \Phi(f(x_2))\bar{\mu}(dx_2) - \Phi\left(\int f(x_2)\,\bar{\mu}(dx_2)\right) \\
&= \int\int \Phi(f(x_2))\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) - \Phi\left(\int\int f(x_2)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1)\right) \\
&\overset{(64)}{=} \int \left(\int \Phi(f(x_2))\,\mu(dx_2|x_1) - \Phi\left(\int f(x_2)\,\mu(dx_2|x_1)\right)\right)\bar{\mu}(dx_1) \\
&\quad + \int \Phi(\bar{f}(x_1))\,\bar{\mu}(dx_1) - \Phi\left(\int \bar{f}(x_1)\,\bar{\mu}(dx_1)\right).
\end{aligned}
$$

According to our assumptions (55) and (56), this yields the estimate

$$
\begin{aligned}
&\left|\int \Phi(f(x_2))\bar{\mu}(dx_2) - \Phi\left(\int f(x_2)\,\bar{\mu}(dx_2)\right)\right| \\
&\leq \int \frac{1}{\rho_2}\int \frac{1}{2f(x_2)}|\nabla_2 f(x_2)|^2\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) \\
&\quad + \frac{1}{\bar{\rho}_1}\int \frac{1}{2\bar{f}(x_1)}|\nabla_1 \bar{f}(x_1)|^2\,\bar{\mu}(dx_1) \\
&= \frac{1}{\rho_2}\int \frac{1}{2f}|\nabla_2 f|^2\,\bar{\mu}(dx_2) + \frac{1}{\bar{\rho}_1}\int \frac{1}{2\bar{f}}|\nabla_1 \bar{f}|^2\,\bar{\mu}(dx_1).
\end{aligned} \tag{65}
$$

We now apply Lemma 6 to the last term. Since f does not depend on x_1, this yields

$$
\begin{aligned}
\int \frac{1}{2\bar{f}}|\nabla_1 \bar{f}|^2\,\bar{\mu}(dx_1) &\leq \frac{\kappa_{12}^2}{\rho_2^2}\int \frac{1}{2f}|\nabla_2 f|^2\,d\mu \\
&= \frac{\kappa_{12}^2}{\rho_2^2}\int \frac{1}{2f}|\nabla_2 f|^2\,\bar{\mu}(dx_2).
\end{aligned} \tag{66}
$$

Inserting (66) into (65) yields as desired

$$
\begin{aligned}
&\left|\int \Phi(f(x_2))\bar{\mu}(dx_2) - \Phi\left(\int f(x_2)\,\bar{\mu}(dx_2)\right)\right| \\
&\qquad\qquad \leq \left(\frac{1}{\rho_2} + \frac{1}{\bar{\rho}_1}\frac{\kappa_{12}^2}{\rho_2^2}\right)\int \frac{1}{2f}|\nabla_2 f|^2\,\bar{\mu}(dx_2).
\end{aligned}
$$

Proof of Corollary 2.

By an approximation argument, we may assume that we have for the marginals

$$\bar\mu(dx_1) \quad \text{satisfies LSI}(\bar\rho_1),$$

$$\bar\mu(dx_2) \quad \text{satisfies LSI}(\bar\rho_2),$$

for some constants $\bar\rho_1, \bar\rho_2 > 0$. Lemma 7 now yields

$$\frac{1}{\bar\rho_2} \;\leq\; \frac{1}{\rho_2} + \frac{1}{\bar\rho_1}\,\frac{\kappa_{12}^2}{\rho_2^2}. \tag{67}$$

By symmetry, we also have

$$\frac{1}{\bar\rho_1} \;\leq\; \frac{1}{\rho_1} + \frac{1}{\bar\rho_2}\,\frac{\kappa_{12}^2}{\rho_1^2}. \tag{68}$$

Inserting (67) into (68) yields

$$\frac{1}{\bar\rho_1} \;\leq\; \frac{1}{\rho_1} + \left(\frac{1}{\rho_2} + \frac{1}{\bar\rho_1}\,\frac{\kappa_{12}^2}{\rho_2^2}\right)\frac{\kappa_{12}^2}{\rho_1^2}$$

and thus

$$\left(1 - \frac{\kappa_{12}^4}{\rho_1^2\,\rho_2^2}\right)\frac{1}{\bar\rho_1} \;\leq\; \frac{1}{\rho_1} + \frac{1}{\rho_2}\,\frac{\kappa_{12}^2}{\rho_1^2},$$

which we rewrite as

$$\left(1 - \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\left(1 + \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\frac{1}{\bar\rho_1} \;\leq\; \left(1 + \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\frac{1}{\rho_1}.$$

This yields

$$\left(1 - \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\frac{1}{\bar\rho_1} \;\leq\; \frac{1}{\rho_1}$$

and thus

$$\bar\rho_1 \;\geq\; \left(1 - \frac{\kappa_{12}^2}{\rho_1\,\rho_2}\right)\rho_1 \;=\; \rho_1 - \frac{\kappa_{12}^2}{\rho_2},$$

which is positive by assumption.

Proof of Lemma 8.

As the starting point we have the two formulas

$$\nabla_1 \bar{H}(x_1,x_2) \;=\; \frac{\int \nabla_1 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\, dx_3}{\int \exp(-H(x_1,x_2,x_3))\, dx_3}$$

and

$$\begin{aligned}
&\nabla_2\nabla_1 \bar{H}(x_1,x_2)\\
&= \frac{\int \nabla_2\nabla_1 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\, dx_3}{\int \exp(-H(x_1,x_2,x_3))\, dx_3}\\
&\quad -\frac{\int \nabla_1 H(x_1,x_2,x_3)\otimes \nabla_2 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\, dx_3}{\int \exp(-H(x_1,x_2,x_3))\, dx_3}\\
&\quad +\frac{\int \nabla_1 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\, dx_3}{\int \exp(-H(x_1,x_2,x_3))\, dx_3}\\
&\quad \otimes\frac{\int \nabla_2 H(x_1,x_2,x_3)\, \exp(-H(x_1,x_2,x_3))\, dx_3}{\int \exp(-H(x_1,x_2,x_3))\, dx_3}\\
&= \int \nabla_2\nabla_1 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2)\\
&\quad -\Big(\int \nabla_1 H(x_1,x_2,x_3)\otimes \nabla_2 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2)\\
&\qquad -\int \nabla_1 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2)\otimes \int \nabla_2 H(x_1,x_2,x_3)\, \mu(dx_3|x_1,x_2)\Big).
\end{aligned}$$

According to Corollary 1 applied to $\mu(dx_3|x_1,x_2)$, $g(x_3) = \nabla_1 H(x_1,x_2,x_3)$, and $h(x_3) = \nabla_2 H(x_1,x_2,x_3)$, we have the inequality

$$\begin{aligned}
|\nabla_2\nabla_1 \bar{H}(x_1,x_2)| &\le \sup_{x_3} |\nabla_2\nabla_1 H(x_1,x_2,x_3)|\\
&\quad + \frac{1}{\rho_3} \sup_{x_3} |\nabla_3\nabla_1 H(x_1,x_2,x_3)| \sup_{x_3} |\nabla_3\nabla_1 H(x_1,x_2,x_3)|\\
&\le \kappa_{12} + \frac{1}{\rho_3}\, \kappa_{13}\, \kappa_{23}.
\end{aligned}$$

Taking the sup over (x_1,x_2) yields (58).

PROOF OF LEMMA 9.

We prove Lemma 9 by induction in N. The case $N = 1$ is trivial. We now assume that the lemma holds for $N-1 \ge 1$ and argue that it also holds for

N. To this purpose, we introduce the related block partitioning of A:

$$A \;=\; \begin{pmatrix} A' & -\kappa \\ \\ -\kappa^t & \rho \end{pmatrix}.$$

The inverse A^{-1} is given by the block partitioning

$$A^{-1} \;=\; \begin{pmatrix} (A')^{-1} + \dfrac{(A')^{-1}\kappa \otimes (A')^{-1}\kappa}{\rho - \kappa \cdot (A')^{-1}\kappa} & \dfrac{(A')^{-1}\kappa}{\rho - \kappa \cdot (A')^{-1}\kappa} \\ \\ \left(\dfrac{(A')^{-1}\kappa}{\rho - \kappa \cdot (A')^{-1}\kappa}\right)^t & \dfrac{1}{\rho - \kappa \cdot (A')^{-1}\kappa} \end{pmatrix}.$$

Based on this representation, we now argue that the entries of A^{-1} are non–negative. As an immediate consequence of the positive definiteness of A we have

$$\rho - \kappa \cdot (A')^{-1}\kappa \;>\; 0, \tag{69}$$

so that

$$(A^{-1})_{NN} \;=\; \frac{1}{\rho - \kappa \cdot (A')^{-1}\kappa} \;\geq\; 0.$$

Combining the induction hypothesis applied to A', i.e.

$$((A')^{-1})_{ij} \;\geq\; 0 \quad \text{for } i, j \in \{1, \cdots, N-1\}, \tag{70}$$

with our assumption

$$\kappa_j \;\geq\; 0 \quad \text{for } j \in \{1, \cdots, N-1\}, \tag{71}$$

we obtain

$$((A')^{-1}\kappa)_i \;=\; \sum_j ((A')^{-1})_{ij}\kappa_j \;\geq\; 0 \quad \text{for } i \in \{1, \cdots, N-1\}, \tag{72}$$

Together with (69), we obtain

$$(A^{-1})_{iN} \;=\; \frac{((A')^{-1}\kappa)_i}{\rho - \kappa \cdot (A')^{-1}\kappa} \;\geq\; 0 \quad \text{for } i \in \{1, \cdots, N-1\}. \tag{73}$$

From (72) we obtain

$$((A')^{-1}\kappa \otimes (A')^{-1}\kappa)_{ij} \;=\; ((A')^{-1}\kappa)_i\,((A')^{-1}\kappa)_j \;\geq\; 0 \quad \text{for } i,j \in \{1,\cdots,N-1\}$$

and together with (69) and (70):

$$(A^{-1})_{ij} \;=\; ((A')^{-1})_{ij} + \frac{((A')^{-1}\kappa \otimes (A')^{-1}\kappa)_{ij}}{\rho - \kappa \cdot (A')^{-1}\kappa} \;\geq\; 0 \quad \text{for } i,j \in \{1,\cdots,N-1\}.$$

Proof of Theorem 1.

We shall prove the seemingly stronger result

$$\begin{aligned} &\forall\, f(x_1,\cdots,x_N) \geq 0 \qquad \int \Phi(f)\,d\mu - \Phi\left(\int f d\mu\right) \\ &\leq \sum_{i,j\in\{1,\cdots,N\}} (A^{-1})_{ij} \left(\int \frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2} \left(\int \frac{1}{2f}|\nabla_j f|^2\,d\mu\right)^{1/2}, \end{aligned} \tag{74}$$

where $(A^{-1})_{ij}$ denote the coefficients of the inverse A^{-1} of A. Statement (74) indeed implies (6): According to (5) we have $A^{-1} \leq \frac{1}{\rho}\,\mathrm{id}$ in the sense of quadratic forms so that (74) implies

$$\begin{aligned} &\int \Phi(f)\,d\mu - \Phi\left(\int f d\mu\right) \\ &\leq \sum_{i,j\in\{1,\cdots,N\}} \frac{1}{\rho}\,\delta_{ij} \left(\int \frac{1}{2f}|\nabla_i f|^2\,d\mu\right)^{1/2} \left(\int \frac{1}{2f}|\nabla_j f|^2\,d\mu\right)^{1/2} \\ &= \frac{1}{\rho} \int \frac{1}{2f} \sum_{i\in\{1,\cdots,N\}} |\nabla_i f|^2\,d\mu. \end{aligned}$$

We show (74) by induction in N. For $N = 1$, the statement (74) is a trivial consequence of our assumption (3). We thus assume that we know (74) for any $(N-1)$–component system and argue that it holds for N. It will be convenient to work with the related block decomposition of A:

$$A \;=\; \begin{pmatrix} A' & -\kappa_N \\ & \\ -\kappa_N^t & \rho_N \end{pmatrix}. \tag{75}$$

Denote by $\bar{A}$ the $(N-1)\times(N-1)$–matrix defined by

$$\bar{A} \;=\; A' - \frac{1}{\rho_N}\,\kappa_N \otimes \kappa_N. \tag{76}$$

We observe that $\bar{A}$ inherits our assumptions on A: It is symmetric and positive definite.

We start by considering the system $\bar{\mu}(dx_1 \cdots dx_{N-1})$, i.e. the marginal of $\mu(dx_1 \cdots dx_N)$ on $X_1 \times \cdots \times X_{N-1}$. Its Hamiltonian is given by

$$\bar{H}(x_1, \cdots, x_{N-1}) \;=\; -\ln \int \exp(-H(x_1, \cdots, x_{N-1}, x_N))\, dx_N.$$

Let $i < j \in \{1, \cdots, N-1\}$ be arbitrary. Lemma 8 applied to $\mu(dx_i dx_j dx_N | \cdots)$ yields

$$\forall\ (x_1, \cdots, x_{N-1}) \quad |\nabla_i \nabla_j \bar{H}(x_1, \cdots, x_{N-1})| \;\le\; \bar{\kappa}_{ij}$$

with

$$\bar{\kappa}_{ij} \;\le\; \kappa_{ij} + \frac{\kappa_{iN}\,\kappa_{jN}}{\rho_N} \;\stackrel{(76)}{=}\; -\bar{A}_{ij}.$$

Now let $i \in \{1, \cdots, N-1\}$ be arbitrary. Corollary 2 applied to $\mu(dx_i dx_N | \cdots)$ yields

$$\begin{aligned} &\forall\ (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{N-1}) \\ &\bar{\mu}(dx_i | x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{N-1}) \quad \text{satisfies LSI}(\bar{\rho}_i) \end{aligned} \tag{77}$$

with

$$\bar{\rho}_i \;\ge\; \rho_i - \frac{\kappa_{iN}^2}{\rho_N} \;\stackrel{(76)}{=}\; \bar{A}_{ii}.$$

Thus, we may apply the induction hypothesis to $\bar{\mu}(dx_1 \cdots dx_{N-1})$ and $\bar{A}$:

$$\forall\, \bar{f}(x_1, \cdots, x_{N-1}) \geq 0 \quad \int \Phi(\bar{f})\, d\bar{\mu} - \Phi\left(\int \bar{f} d\bar{\mu}\right)$$
$$\leq \sum_{i,j \in \{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij} \left(\int \frac{1}{2\bar{f}} |\nabla_i \bar{f}|^2\, d\bar{\mu}\right)^{1/2} \left(\int \frac{1}{2\bar{f}} |\nabla_j \bar{f}|^2\, d\bar{\mu}\right)^{1/2}. \tag{78}$$

Now let $f(x_1, \cdots, x_N) \geq 0$ be given and set

$$\bar{f}(x_1, \cdots, x_{N-1}) \;:=\; \int f(x_1, \ldots, x_{N-1}, x_N)\, \mu(dx_N | x_1, \cdots, x_{N-1}).$$

As in the proof of Lemma 7, we split the left–hand side of (74):

$$\int \Phi(f)\, d\mu - \Phi\left(\int f\, d\mu\right)$$
$$= \int \left(\int \Phi(f)\, \mu(dx_N|\cdots) - \Phi\left(\int f\, \mu(dx_N|\cdots)\right)\right)$$
$$\bar{\mu}(dx_1 \cdots dx_{N-1})$$
$$+ \int \Phi(\bar{f})\, d\bar{\mu} - \Phi\left(\int \bar{f}\, d\bar{\mu}\right). \tag{79}$$

By assumption (3) we have

$$\int \Phi(f)\, \mu(dx_N|\cdots) - \Phi\left(\int f\, \mu(dx_N|\cdots)\right)$$
$$\leq \frac{1}{\rho_N} \int \frac{1}{2f} |\nabla_N f|^2\, \mu(dx_N|\cdots),$$

so that we obtain for the first right–hand side term in (79):

$$\int \left(\int \Phi(f)\, \mu(dx_N|\cdots) - \Phi\left(\int f\, \mu(dx_N|\cdots)\right)\right) \bar{\mu}(dx_1 \cdots dx_{N-1})$$
$$\leq \frac{1}{\rho_N} \int \frac{1}{2f} |\nabla_N f|^2\, d\mu. \tag{80}$$

We apply (78) to the second right–hand side term in (79). Combining this with (80), (79) becomes

$$
\begin{aligned}
&\int \Phi(f)\, d\mu - \Phi\left(\int f d\mu\right) \\
&\leq \sum_{i,j\in\{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij} \left(\int \frac{1}{2\bar{f}}|\nabla_i \bar{f}|^2\, d\bar{\mu}\right)^{1/2} \left(\int \frac{1}{2\bar{f}}|\nabla_j \bar{f}|^2\, d\bar{\mu}\right)^{1/2} \\
&\quad + \frac{1}{\rho_N} \int \frac{1}{2f}|\nabla_N f|^2\, d\mu.
\end{aligned} \tag{81}
$$

We now want to express the first right–hand side terms of (81) in terms of f and μ. To this purpose, let $i \in \{1, \cdots, N-1\}$ be arbitrary. Because of our assumptions (2) and (3), we may apply Lemma 6 to $\mu(dx_i dx_N|\cdots)$ and obtain

$$
\begin{aligned}
&\left(\int \frac{1}{2\bar{f}}|\nabla_i \bar{f}|^2\, \bar{\mu}(dx_i|\cdots)\right)^{1/2} \\
&\leq \left(\int \frac{1}{2f}|\nabla_i f|^2\, \mu(dx_i dx_N|\cdots)\right)^{1/2} + \frac{\kappa_{iN}}{\rho_N} \left(\int \frac{1}{2f}|\nabla_N f|^2\, \mu(dx_i dx_N|\cdots)\right)^{1/2}.
\end{aligned}
$$

By the triangle inequality in $L^2(\bar{\mu}(dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_{N-1}))$, this yields

$$
\begin{aligned}
&\left(\int \frac{1}{2\bar{f}}|\nabla_i \bar{f}|^2\, d\bar{\mu}\right)^{1/2} \\
&\leq \left(\int \frac{1}{2f}|\nabla_i f|^2\, d\mu\right)^{1/2} + \frac{\kappa_{iN}}{\rho_N} \left(\int \frac{1}{2f}|\nabla_N f|^2\, d\mu\right)^{1/2}.
\end{aligned} \tag{82}
$$

Since $\bar{A}$ has non–positive off–diagonal entries, an application of Lemma 9 yields that all entries of $(\bar{A})^{-1}$ are non–negative. Thus we may insert the inequality (82) into (81):

$$
\begin{aligned}
&\int \Phi(f)\, d\mu - \Phi\left(\int f d\mu\right) \\
&\le \sum_{i,j\in\{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij} \left(\int \frac{1}{2f}|\nabla_i f|^2\, d\mu\right)^{1/2} \left(\int \frac{1}{2f}|\nabla_j f|^2\, d\mu\right)^{1/2} \\
&\quad +2 \sum_{i,j\in\{1,\cdots,N-1\}} (\bar{A}^{-1})_{ij} \frac{\kappa_{jN}}{\rho_N} \left(\int \frac{1}{2f}|\nabla_i f|^2\, d\mu\right)^{1/2} \left(\int \frac{1}{2f}|\nabla_N f|^2\, d\mu\right)^{1/2} \\
&\quad + \left(\frac{1}{\rho_N} + \sum_{i,j\in\{1,\cdots,N-1\}} \frac{\kappa_{iN}}{\rho_N} (\bar{A}^{-1})_{ij} \frac{\kappa_{jN}}{\rho_N}\right) \int \frac{1}{2f}|\nabla_N f|^2\, d\mu. \qquad (83)
\end{aligned}
$$

We now argue that (83) and (74) coincide. We recall the block partitioning of A^{-1}:

$$
A^{-1} = \begin{pmatrix} (A')^{-1} + \dfrac{(A')^{-1}\kappa_N \otimes (A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} & \dfrac{(A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} \\[2ex] \left(\dfrac{(A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N}\right)^t & \dfrac{1}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} \end{pmatrix}.
$$

In view of (76), the statement reduces to the following identities:

$$
(A')^{-1} + \frac{(A')^{-1}\kappa_N \otimes (A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} = \left(A' - \frac{1}{\rho_N}\kappa_N \otimes \kappa_N\right)^{-1}, \qquad (84)
$$

$$
\frac{(A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} = \left(A' - \frac{1}{\rho_N}\kappa_N \otimes \kappa_N\right)^{-1} \frac{\kappa_N}{\rho_N}, \qquad (85)
$$

$$
\frac{1}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} = \frac{1}{\rho_N} + \frac{\kappa_N}{\rho_N} \cdot \left(A' - \frac{1}{\rho_N}\kappa_N \otimes \kappa_N\right)^{-1} \frac{\kappa_N}{\rho_N}. \qquad (86)
$$

Identity (84) is seen to hold as follows:

$$
\begin{aligned}
&\left(A' - \frac{1}{\rho_N}\kappa_N \otimes \kappa_N\right) \left((A')^{-1} + \frac{(A')^{-1}\kappa_N \otimes (A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N}\right) \\
&= \mathrm{id}' + \left(\frac{1}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} - \frac{1}{\rho_N} - \frac{\kappa_N \cdot (A')^{-1}\kappa_N}{\rho_N\,(\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N)}\right) \\
&\qquad \cdot \kappa_N \otimes (A')^{-1}\kappa_N.
\end{aligned}
$$

The prefactor in the parenthesis vanishes. Identity (85) is a consequence of

(84):

$$
\begin{aligned}
&\left(A' - \frac{1}{\rho_N}\kappa_N \otimes \kappa_N\right)^{-1} \frac{\kappa_N}{\rho_N} \\
&\overset{(84)}{=} \left((A')^{-1} + \frac{(A')^{-1}\kappa_N \otimes (A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N}\right) \frac{\kappa_N}{\rho_N} \\
&= \frac{1}{\rho_N}\left(1 + \frac{\kappa_N \cdot (A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N}\right)(A')^{-1}\kappa_N \\
&= \frac{1}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N}(A')^{-1}\kappa_N.
\end{aligned}
$$

Finally, identity (86) is a consequence of (85):

$$
\begin{aligned}
&\frac{1}{\rho_N} + \frac{\kappa_N}{\rho_N} \cdot \left(A' - \frac{1}{\rho_N}\kappa_N \otimes \kappa_N\right)^{-1} \frac{\kappa_N}{\rho_N} \\
&\overset{(85)}{=} \frac{1}{\rho_N} + \frac{1}{\rho_N}\frac{\kappa_N \cdot (A')^{-1}\kappa_N}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N} \\
&= \frac{1}{\rho_N - \kappa_N \cdot (A')^{-1}\kappa_N}.
\end{aligned}
$$

Proof of Theorem 2.

Recalling the definition of $\bar{f}$ from (54), we integrate the identity

$$
\Phi(f) \;=\; f\,\log\bar{f} \;+\; \bar{f}\,\Phi\Big(\frac{f}{\bar{f}}\Big)
$$

and use Definition 2 to obtain the usual decomposition of entropy:

$$
\begin{aligned}
&\int_X \Phi(f)d\mu \\
&= \int_{X_1}\int_{X_2} f\,\log\bar{f}\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) + \int_{X_1}\int_{X_2} \bar{f}\,\Phi\Big(\frac{f}{\bar{f}}\Big)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) \\
&= \int_{X_1}\int_{X_2} f\,\mu(dx_2|x_1)\,\log\bar{f}\,\bar{\mu}(dx_1) + \int_{X_1}\bar{f}\int_{X_2}\Phi\Big(\frac{f}{\bar{f}}\Big)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1) \\
&\overset{(27)}{=} \int_{X_1}\Phi(\bar{f})\,\bar{\mu}(dx_1) \;+\; \int_{X_1}\bar{f}\int_{X_2}\Phi\Big(\frac{f}{\bar{f}}\Big)\,\mu(dx_2|x_1)\,\bar{\mu}(dx_1). \qquad (87)
\end{aligned}
$$

We start with the second term on the right–hand side of (87). According to

(22), we have for all $x_1 \in X_1$

$$\begin{aligned}\int_{X_2} \Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1) \;&\le\; \Phi\Big(\int_{X_2} \frac{f}{\bar f}\,\mu(dx_2|x_1)\Big) + \frac{1}{\rho_2}\int_{X_2} \frac{\bar f}{2f}\Big|\nabla_2 \frac{f}{\bar f}\Big|^2 \mu(dx_2|x_1)\\ &\overset{(27)}{=}\; \frac{1}{\rho_2}\frac{1}{\bar f}\int_{X_2}\frac{1}{2f}\,|\nabla_2 f|^2\,\mu(dx_2|x_1),\end{aligned}$$

so that by integrating we obtain

$$\begin{aligned}\int_{X_1} \bar f \int_{X_2} \Phi\Big(\frac{f}{\bar f}\Big)\,\mu(dx_2|x_1)\,\bar\mu(dx_1) \;&\le\; \frac{1}{\rho_2}\int_{X_1}\int_{X_2} \frac{1}{2f}|\nabla_2 f|^2\,\mu(dx_2|x_1)\,\bar\mu(dx_1)\\ &=\; \frac{1}{\rho_2}\int_X \frac{1}{2f}|\nabla_2 f|^2\,d\mu. \qquad (88)\end{aligned}$$

We now turn to the first term on the right–hand side of (87). According to (23), we have

$$\begin{aligned}\int_{X_1} \Phi(\bar f)\,\bar\mu(dx_1) \;&\le\; \Phi\Big(\int_{X_1} \bar f\,\bar\mu(dx_1)\Big) + \frac{1}{\bar\rho_1}\int_{X_1}\frac{1}{2\bar f}|\nabla_1 \bar f|^2\,\bar\mu(dx_1)\\ &\overset{(27)}{=}\; \Phi\Big(\int_X f\,d\mu\Big) + \frac{1}{\bar\rho_1}\int_{X_2}\frac{1}{2\bar f}|\nabla_1 \bar f|^2\,\bar\mu(dx_1). \qquad (89)\end{aligned}$$

By Lemma 6 and Young's inequality, we have for any $\tau \in (0,1)$,

$$\begin{aligned}&\int \frac{1}{2\bar f}|\nabla_1 \bar f|^2\,\bar\mu(dx_1)\\ &\le\; \frac{1}{\tau}\int \frac{1}{2f}|\nabla_1 f|^2\,d\mu \;+\; \frac{1}{1-\tau}\frac{\bar\rho_1\alpha}{\rho_2}\int \frac{1}{2f}|\nabla_2 f|^2\,d\mu,\end{aligned}$$

so that (89) becomes

$$\begin{aligned}&\int_{X_1} \Phi(\bar f)\,\bar\mu(dx_1)\\ &\overset{(24)}{\le}\; \Phi\Big(\int_X f\,d\mu\Big) + \frac{1}{\tau\bar\rho_1}\int \frac{1}{2f}|\nabla_1 f|^2\,d\mu \;+\; \frac{1}{1-\tau}\frac{\alpha}{\rho_2}\int \frac{1}{2f}|\nabla_2 f|^2\,d\mu. \qquad (90)\end{aligned}$$

Substituting (88) and (90) into (87) gives:

$$\int_X \Phi(f)d\mu \;\le\; \Phi\Big(\int_X f\,d\mu\Big) + \frac{1}{\rho_2}\int_X \frac{1}{2f}|\nabla_2 f|^2\,d\mu$$
$$+\frac{1}{\tau}\frac{1}{\bar\rho_1}\int_X \frac{1}{2f}|\nabla_1 f|^2\,d\mu + \frac{1}{1-\tau}\frac{\alpha}{\rho_2}\int_X \frac{1}{2f}|\nabla_2 f|^2 d\mu.$$

Since $|\nabla f|^2 = |\nabla_1 f|^2 + |\nabla_2 f|^2$, this yields the bound on the LSI constant:

$$\frac{1}{\rho} \;\le\; \max\left\{\frac{1}{\tau}\frac{1}{\bar\rho_1}, \frac{1}{\rho_2} + \frac{1}{1-\tau}\frac{\alpha}{\rho_2}\right\}. \tag{91}$$

The optimization in τ completes the proof. Indeed, the optimal τ in (91) is characterized by

$$\frac{1}{\tau\bar\rho_1} \;=\; \frac{1}{\rho_2} + \frac{\alpha}{(1-\tau)\rho_2},$$

that is,

$$(1-\tau)\rho_2 \;=\; \bar\rho_1\tau(1-\tau) + \alpha\bar\rho_1\tau.$$

The admissible solution is

$$\tau \;=\; \frac{1}{2}\Big((1+\alpha) + \frac{\rho_2}{\bar\rho_1} - \sqrt{(1+\alpha+\frac{\rho_2}{\bar\rho_1})^2 - 4\frac{\rho_2}{\bar\rho_1}}\Big),$$

so that (91) turns as desired into

$$\rho \ge \frac{1}{2}\Big(\rho_2 + (1+\alpha)\bar\rho_1 - \sqrt{(\rho_2 + (1+\alpha)\bar\rho_1)^2 - 4\,\bar\rho_1\,\rho_2}\Big).$$

Acknowledgements

We thank G. Blower, F. Bolley, A. Eberle, M. Rost, and H.T. Yau for inspiring discussions. We also gratefully acknowledge N. Grunewald and C. Villani, our co-authors on [GORV]. F. Otto was partially supported by the DFG through SFB 611. M.G. Reznikoff was partially supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

References

[BE] D. Bakry and M. Emery, Diffusions Hypercontractives, *Sem. Probab. XIX, Lecture Notes in Math.*, **1123**, Springer–Verlag, (1985), pp. 177–206.

[BB] G. Blower and F. Bolley, Concentration of Measure on Product Spaces with Applications to Markov Processes, submitted for publication.

[BH] T. Bodineau and B. Hellfer, The Log–Sobolev Inequality for Unbounded Spin Systems, *Journal of Functional Analysis*, **166**, (1999), pp. 168–178.

[G] L. Gross, Logarithmic Sobolev Inequalities, *Amer. J. Math.*, **97**, (1975), pp. 1061–1083.

[GORV] N. Grunewald, F. Otto, M. G. Reznikoff, and C. Villani, A Two–Scale Proof of a Logarithmic Sobolev Inequality, SFB preprint 2005.

[GZ] A. Guionnet and B. Zegarlinski, Lectures on Logarithmic Sobolev Inequalities, *Séminaire de Probabilités, XXXVI, Lecture Notes in Math.*, **1801**, Springer–Verlag, (2003), pp. 1–134.

[HS] R. Holley and D.Stroock, Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Statist. Phys.*, **46**, (1987), pp. 1159–1194.

[L] M. Ledoux, Logarithmic Sobolev Inequalities for Unbounded Spin Systems Revisited, *Séminaire de Probabilités, XXXV, Lecture Notes in Math.*, **1755**, Springer–Verlag (2001), pp. 167–194.

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

[O] F. Otto, The Geometry of Dissipative Evolution Equations: The Porous Medium Equation, *Commun. Partial Differential Equations*, **26**, (2001), pp. 101–174.

[OV] F. Otto and C. Villani, Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality, *J. Funct. Analysis*, **173**, (2000), pp. 361–400.

[R] G. Royer, Une Initiation aux Inégalités de Sobolev Logarithmiques, Cours Spécialisés, Soc. Math. de France, (1999).

[SZ] D. Stroock and B. Zegarlinski, The Logarithmic Sobolev Inequality for Continuous Spin Systems on a Lattice, *J. Func. Anal.*, **104**, (1992), pp. 299–326.

[Yo] N. Yoshida, The Log–Sobolev Inequality for Weakly Coupled Lattice Fields, *Probab. Theor. Relat. Fields*, **115**, (1999), pp. 1–40.

[Yo2] N. Yoshida, The Equivalence of the Log–Sobolev Inequality and a Mixing Condition for Unbounded Spin Systems on the Lattice, *Ann. Inst. H. Poincaré Probab. Statist.*, **37**, (2001), pp. 223–243.

[Z2] B. Zegarlinski, Log–Sobolev Inequalities for Infinite One Dimensional Lattice Systems, *Commun. Math. Phys.*, **133**, (1990), pp. 147–162.

[Z1] B. Zegarlinski, On Log–Sobolev Inequalities for Infinite Lattice Systems, *Lett. Math. Phys.*, **20**, (1990), pp. 173–182.

[Z3] B. Zegarlinski, The Strong Decay to Equilibrium for the Stochastic Dynamics of Unbounded Spin Systems on a Lattice, *Comm. Math. Phys.*, **175**, (1996), pp. 401–432.