Spectral Gap and Logarithmic Sobolev Inequality for Kawasaki and Glauber Dynamics

Sheng Lin Lu[1], Horng-Tzer Yau[2,*]

[1] Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1092, USA
e-mail: shenglu@math.lsa.umich.edu
[2] Courant Institute, New York University, New York, NY 10012, USA
e-mail: yau@math.nyu.edu

Abstract. We prove that the spectral gap of the Kawasaki dynamics shrink at the rate of $1/L^2$ for cubes of size L provided that some mixing conditions are satisfied. We also prove that the logarithmic Sobolev inequality for the Glauber dynamics in standard cubes holds uniformly in the size of the cube if the Dobrushin-Shlosman mixing condition holds for standard cubes.

Introduction

As the simplest model in statistical mechanics, Ising model has been studied extensively. It is by far the most studied model in mathematical physics and its phase structures were analyzed in great detail. The study of dynamical properties of the Ising model, on the other hand, is in a much more primitive stage. Our main concern is the hydrodynamical limit of the Ising model for which we shall provide a basic estimate on the gap of Kawasaki dynamics. The hydrodynamical limit of various models has been studied recently and several useful methods were developed, see, e.g. [DP, S] for a review. A central assumption of these methods is the so-called *gradient condition*. Roughly speaking, it means that the current of the dynamics is by itself a gradient of some other quantity. For models with this property, a natural summation by parts can be performed and the technical difficulty is greatly reduced. The drawback of gradient models is that the diffusion coefficient, as given by the Green-Kubo formula, is determined by the thermodynamical quantities rather than depending on correlation functions as the nongradient model does. Therefore, it does not manifest effects of fluctuations on the diffusion coefficient.

Another interesting aspect of the gradient condition is that, except in dimension $d = 1$ or the infinite temperature case, no gradient model has been constructed for any truly interacting, reversible models with discrete spin space. So a study of the

nongradient model with discrete spin space is an essential step if any progress on the hydrodynamical limit for the Ising model can be made. So far the only known result of the nongradient system is Varadhan's work [V, see also KLO, Q] on Ginzburg-Landau type dynamics with product invariant measure. One of the key estimates needed is a bound on the spectral gap which in the special case of product measure can be proved rather straightforwardly. In this paper we shall prove such a gap is true for Ising type models under certain mixing assumptions on the Gibbs state. Let us sketch our results briefly here.

Let Λ be a cube in $\mathbb{Z}^d$ of size L and let μ_ω denote a Gibbs state with boundary condition ω. Since our dynamics conserves the total number of particles, it is natural to introduce the "canonical Gibbs state" $\nu_{\omega,N}$ with the total number of particles (or up spins) fixed. Let A denote the standard Kawasaki dynamics (with Dirichlet form given by (1.23) below) with reversible measure $\nu_{\omega,N}$. Our main result is that the spectral gap of A cannot shrink faster than $1/L^2$ if some mixing conditions are satisfied, see assumptions A1–A3 in Sect. 1. An upper bound on the gap of the order $1/L^2$ can be easily obtained by considering a slowly varying test function. Thus this correctly pins down the decay rate of the gap.

Our methods are based on the martingale approach. It also proves that there is a positive spectral gap for the Glauber dynamics (with Dirichlet form given by (1.21) below) uniformly with respect to the volume and boundary conditions if certain mixing conditions are satisfied (see Assumption (A.1) in Sect. 1). With the method almost unchanged, a logarithmic Sobolev inequality is also proved under the same assumption. It should be emphasized that these results are general in the sense that they applied to any models with finite range interactions (or summable interactions) with discrete or continuous spins for which assumption A.1 holds. If one is interested only in ferromagnetic Ising models, a useful tool known as attractiveness becomes available and stronger results can be obtained. Recently Martinelli and Olivieri [MO1] have proved the important result that exponential convergence holds for ferromagnetic Ising models up to the critical temperature. For general models, they also obtained results similar to ours (Theorem 1 and 3) independently with different arguments [MO2] (see also the next paragraph for a comparison with [SZ]). Although the mixing conditions (A.1) assumed here are equivalent to theirs [O, OP], their proof has the advantage of being directly based on mixing conditions for only one cube. For the Kawasaki dynamics, we are not aware of any result except in the case of independent random variables (with the global constraint that the total magnetizations is fixed) [F, KLO, Q].

Let us pause to comment on some history of the spectral gap and logarithmic Sobolev inequality for the Gibbs states. The importance of the logarithmic Sobolev inequality and its connection to the hypercontractivity (for general measures) was first proved by L. Gross in his 1976 paper [L]. (See e.g. [DGS] for a review.) Since then it has been used as an important tool to understand the exponential convergence to equilibrium. For Glauber dynamics, there are extensive literatures on this subject since the late seventies by, e.g., Holley, Liggett, Stroock et al. Most of these results are one dimensional or concern some general properties (e.g. [CS]). A higher dimensional result was obtained by Aizenman and Holley [AH] which states that the spectral gap for the infinite volume Glauber dynamics is strictly positive if the Dobrushin-Shlosman uniqueness condition is satisfied. Later on Zegarlinski [Z2] proved the logarithmic Sobolev inequality under the Dobrushin uniqueness condition. Recently Stroock and Zegarlinski [SZ] proved that the logarithmic Sobolev inequality is equivalent to the

"Dobrushin-Shlosman mixing conditions" (see also [S, Z3] for a review), which are closely related to the mixing conditions considered in this paper. The alert reader may have found the above mixing or uniqueness conditions confusing. Unfortunately, a closer examination of the literature can only add to the confusion. We shall not discuss the relations among these and other equivalent conditions in this paper except the following remark concerning the comparison between our results in the Glauber dynamics case and that of [SZ]. The interested reader is referred to the recent papers by Martinelli and Olivieri [MO1, MO2] and references therein for a thorough study and clear review of these mixing conditions.

Apart from the fact that we are using a different approach, our result differs from [SZ] in the following way: In [SZ] the DS mixing condition is assumed for all domains in $\mathbb{Z}^d$ and the logarithmic Sobolev inequality is proved for all domains in $\mathbb{Z}^d$ while we assume mixing conditions for standard cubes and prove that the logarithmic Sobolev inequality holds uniformly for all standard cubes. (If one is interested in infinite volume Gibbs states rather than finite volume Gibbs states, the approach of [SZ] also requires only the DS mixing conditions for cubes [Z1].) It was emphasized in [MO1, MO2] that the DS mixing condition for general domains is not expected to hold in low temperature with magnetic field. For example, a two dimensional "cube" in R^3 in low temperature with two boundaries consisting of translates of the two dimensional "cube" will not satisfy the DS condition if the two boundary conditions, say, take value plus one while the magnetic field is minus two so that the effect of magnetic field is completely cancelled by the boundary condition. On the other hand, we do not require the mixing condition for domains other than standard cubes of size $L \times L \times \cdots \times L$. The mixing conditions for cubes rather than for arbitrary domains was emphasized by Olivieri [O] and Olivieri-Picco [OP] in their study of cluster expansion for spin systems. We thank Martinelli and Olivieri for informing us of the importance of assuming mixing conditions only for cubes, the previous comparison between the mixing conditions of [SZ] and ours, and for providing us the previous example.

Unfortunately, so far we are not able to prove the logarithmic Sobolev inequality for the Kawasaki dynamics for the Ising model except for $d = 1$. If one replaces the Ising model by Ginzburg-Landau models then the corresponding logarithmic Sobolev inequality can be proved. It is interesting to note that for the hydrodynamical limit the Ising model is by far the hardest. We shall delay the proof of the logarithmic Sobolev inequality for Ginzburg-Landau models in a forthcoming paper in the hope that the difficulty with the Ising model can be resolved.

Finally, we comment on the difference between the Kawasaki and Glauber dynamics. In Glauber dynamics, the convergence to equilibrium is exponentially fast and the influence of both the dynamics and the Ising measure itself exponentially decays with the distance. The Kawasaki dynamics, however, does not converge to equilibrium with exponential rate. Furthermore, due to the global condition that the total number of particles is conserved, the canonical Gibbs state is *negatively* correlated, in the sense that $\langle \eta_x; \eta_y \rangle \sim -1/L^d$ for $|x - y| \sim L$ in a cube of size L. (To see this, consider the special case that the canonical Gibbs state degenerates into independent random variables with the constraint $\sum_x \eta_x = \text{const}$. Clearly one has

$$\left\langle \sum_x \eta_x; \sum_x \eta_x \right\rangle = 0 \text{ which implies that}$$

$$\langle \eta_x; \eta_y \rangle = -\langle \eta_x; \eta_x \rangle / L^d \sim -1/L^d).$$

While $1/L^d$ is very small for L large, these negative correlations sum up to order one and in some sense are responsible for the $1/L^2$ decay of the spectral gap. In other words, long range negative correlations play a very significant role in the Kawasaki dynamics while exponential decay dominates in Glauber dynamics. One can easily understand this by considering the infinite temperature case, i.e., the product measure case. While the spectral gap is trivial for the Glauber dynamics, it already requires nontrivial arguments for the Kawasaki dynamics [F, KLO, Q] especially when more than one particle is allowed per lattice site [KLO].

    This paper is organized as follows: Chapter 1 is the statement of main results; in Chapter 2 we prove the spectral gap for Glauber dynamics; Chapters 3 and 4 contain the main technical estimates of the paper and the spectral gap of the Kawasaki dynamics is proved in Chapter 4; Chapter 5 provides some details on the equivalence of ensembles needed in Chapters 3 and 4; finally we prove the logarithmic Sobolev inequality for Glauber dynamics in Chapter 6. For readers interested only in Glauber dynamics, Chapters 3 to 5 can be omitted.

I. Statement of Main Results

Let Λ be a domain in $\mathbb{Z}^d$ and let $\partial\Lambda$ denote its boundary

$$\partial\Lambda = \{y \in \mathbb{Z}^d \setminus \Lambda \mid \text{dist}(y, \Lambda) = 1\}, \tag{1.1}$$

where the distance function is defined by

$$\text{dist}(y, \Lambda) = \inf_{x \in \Lambda} |x - y|,$$
$$|x - y| = \max_{\alpha=1,\ldots,d} |x^\alpha - y^\alpha|. \tag{1.2}$$

Let ω be a configuration on $\partial\Lambda$ where ω_x belongs to some state space X for all $x \in \partial\Lambda$. For simplicity, we shall restrict the state space to be $\mathbb{Z}_2 = \{0, 1\}$. All results in this paper hold if one replaces $\mathbb{Z}_2$ by

$$Z_p = \{0, 1, 2, \ldots, p-1\}, \quad 2 \le p \in N. \tag{1.3}$$

    We shall consider the spectral gap problem in a class of domains which we shall call generalized cubes. Recall the standard cube in $\mathbb{Z}^d$ is characterized by its size L with

$$\Lambda_L = \{x = (x^1, \ldots, x^d) \mid x^i \in \mathbb{Z}^d, \ 1 \le x^i \le L\}. \tag{1.4}$$

By definition, a simple cube is a translation of the standard cube. The boundary $\partial\Lambda_L$ of a simple cube is a union of faces which are cubes in $\mathbb{Z}^{d-1}$. Denote the faces by $\partial^1\Lambda_L, \partial^2\Lambda_L, \ldots, \partial^{2d}\Lambda_L$. We now define the notion of generalized cubes. Choose a lexicographic order in $\mathbb{Z}^\alpha$. Let $F^i \subset \partial^i\Lambda_L$ be a subset of $\partial^i\Lambda_L$ defined by

$$F^i = \{x \in \partial^i\Lambda_L \mid x \ge x_i \in \partial^i\Lambda_L \text{ in } \mathbb{Z}^{d-1}\}, \tag{1.5}$$

where x_i is some fixed point in $\partial^i\Lambda_L$. A generalized cube of size $L+1$ is the union $\Lambda_L \cup F^1 \cup \cdots \cup F^{2d}$. We shall call a generalized cube simply a cube.

    The Hamiltonian we are interested in is the class consisting of translationally invariant, finite range interactions. For simplicity of notation, we shall restrict

ourselves to nearest neighbor interactions. Thus a Hamiltonian is characterized by an interaction $J(\eta_x, \eta_y)$ with

$$H_\Lambda = \sum_{\substack{x,y \in \Lambda \\ |x-y|=1}} J(\eta_x, \eta_y).$$

More generally,

$$H_{\Lambda,\omega,\lambda}(\eta) = \sum_{\substack{x,y \in \Lambda, \\ |x-y|=1}} J(\eta_x, \eta_y) + \sum_{\substack{y \in \partial\Lambda, \ x \in \Lambda, \\ |x-y|=1}} J(\eta_x, \omega_y) + \lambda \sum_{x \in \Lambda} \eta_x. \qquad (1.6)$$

The standard Gibbs state with chemical potential λ and boundary condition ω is characterized by the density

$$d\mu_{\Lambda,\omega,\lambda}(\eta) = \exp[-H(\eta)]/Z_{\Lambda,\omega,\lambda}. \qquad (1.7)$$

Here the partition function $Z_{\Lambda,\omega,\lambda}$ is the normalization factor to make $d\mu_{\Lambda,\omega,\lambda}$ into a probability density. We shall denote by $E^{\mu_{\Lambda,\omega,\lambda}}$ or $E_{\Lambda,\omega,\lambda}$ or $\langle \ \rangle_{\Lambda,\omega,\lambda}$ the expectation with respect to $d\mu_{\Lambda,\omega,\lambda}$. Recall the pressure defined by

$$p_{\Lambda,\omega}(\lambda) = |\Lambda|^{-1} \log Z_{\Lambda,\omega,\lambda}. \qquad (1.8)$$

The infinite volume limit of $p_{\Lambda,\omega}(\lambda)$ exists and is independent of ω, i.e. $\lim_{\Lambda \to \infty} p_{\Lambda,\omega}(\lambda) = p(\lambda)$.

We need the concept of canonical Gibbs states. Let N be a fixed positive integer. Then a canonical Gibbs state with total number of particles N and the boundary condition ω is characterized by the density

$$d\nu_{\Lambda,\omega,N} = d\mu_{\Lambda,\omega} |_{\bar\eta=N}. \qquad (1.9)$$

Here $\bar\eta = \sum_{x \in \Lambda} \eta_x$. Note that the right side of (1.9) is independent of λ since $\bar\eta$ is fixed.
Define the canonical partition function

$$Z^c_{\Lambda,\omega,N} = \sum_{\bar\eta=N} \exp[-H_{\Lambda,\omega}(\eta)], \qquad (1.10)$$

where $H_{\Lambda,\omega} = H_{\Lambda,\omega,\lambda=0}$. *We shall follow the convention to omit the subindex in case it is zero, understood, or unimportant.* Recall the free energy

$$f_{\Lambda,\omega,N} = -|\Lambda|^{-1} \log Z^c_{\Lambda,\omega,N}. \qquad (1.11)$$

The infinite volume limit of f exists if $N/|\Lambda| \to \varrho$ in the limit. Furthermore it is related to the pressure by the Legendre transform

$$f(\varrho) = \sup_\lambda (\lambda\varrho - p(\lambda)). \qquad (1.12)$$

For any function g of the configuration space, define two operators

$$\sigma_x g(\eta) = g(\sigma_x \eta) - g(\eta), \qquad (1.13)$$
$$T_{xy} g(\eta) = g(T_{xy}\eta) - g(\eta). \qquad (1.14)$$

Here $\sigma_x \eta$ and $T_{xy} \eta$ are defined by

$$(\sigma_x \eta)_y = (\eta^x)_y = \delta_{xy}(1 - \eta_y) + (1 - \delta_{xy})\eta_y \,, \qquad (1.15)$$
$$(T_{xy}\eta)_z = (\eta^{xy})_z = \delta_{xz}\eta_y + \delta_{yz}\eta_x + (1 - \delta_{xz} - \delta_{yz})\eta_z \,. \qquad (1.16)$$

To state our main result, we need the following assumptions. Define first the set

$$A_{\omega_1,\omega_2} = \{x \in \partial \Lambda | \omega_1(x) \neq \omega_2(x)\}.$$

Assumption A1. *Let g be a function depending only on the configuration of a subset U in a cube Λ. Then*

$$|E_{\Lambda,\omega_1,\lambda}[g] - E_{\Lambda,\omega_2,\lambda}[g]|$$
$$\leq \text{const. } [\varrho(1 - \varrho)]^{1/2} |A_{\omega_1,\omega_2}| \, |U| \exp[-\text{const. dist}(A_{\omega_1,\omega_2}, U)] \, \|g\|_\infty \,. \quad (1.17)$$

Here $E_{\Lambda,\omega,\lambda}$ denote the expectation with respect to $d\mu_{\Lambda,\omega,\lambda}$ and the constants are independent of Λ and ω_i.

Assumption A2. *There exists a summable function $t(s) \leq \text{const. } s^{-d}$ so that for any local functions f and g with f (g resp.) depending only on configurations in U (V resp.) we have*

$$|E^{\nu_{\Lambda,N,\omega}}[f; g]| \leq C(\varrho)\{L^{-d} + t(\text{dist}(U, V))\} \|f\|_\infty \|g\|_\infty |U| \, |V|, \qquad (1.18)$$

where $C(\varrho) \to 0$ as $\varrho \to 0$. Here Λ is a cube of size L.

Assumption A3. *Let $f = |\Lambda|^{-1} \sum_x f_x$, $g = |\Lambda|^{-1}\sum_x g_x$ and $h = h_y$ with f_x, g_x and h_y being bounded local functions at x and y. Then*

$$E^{\nu_{\Lambda,N,\omega}}[f; g; h] \leq \text{const. } L^{-2d} \|f\|_\infty \|g\|_\infty \|h\|_\infty \,. \qquad (1.19)$$

Here $\langle f; g; h \rangle = \langle (f - \langle f \rangle)(g - \langle g \rangle)(h - \langle h \rangle) \rangle$, $\|f\|_\infty = \sup_x \|f_x\|_\infty$ and Λ is a cube of size L.

Note that in principle the volumes of the supports of f_x, g_x and h_y should appear in (1.19) as in (1.17) and (1.18). We neglect them because all local functions considered in this paper depend only on configurations in cubes of uniformly bounded volumes. We shall adopt the convention that, by "local functions," we mean functions depending only on configurations on a cube of size less than $4dR + 1$ with R denoting the range of the interactions in the Hamiltonian. In particular $R = 1$ for the Ising model.

We shall assume that the domain Λ in Assumption A2 is of the form

$$\Lambda = \Omega \setminus \Gamma$$

with Ω and Γ being generalized cubes and that $|\Gamma| < |\Omega|^\varepsilon$ for some $\varepsilon > 0$, say $\varepsilon = 1/100$. We have assumed Assumption A1 for all cubes. In fact, it can be proved that if (1.17) holds for a fixed cube then it holds for all cubes, [O, OP]. Furthermore, the exponential decay assumption for that fixed cube can be considerably weakened [O, OP]. In any event, we do not exponential decay for Theorem 1 or 3. A power law decay faster than summable will be enough, for example.

Assumption A.1 is a standard assumption in the study of Glauber dynamics and has been studied and reviewed extensively [MO1, St, SZ]. Assumptions A.2 and A.3 are not as familiar and we are not aware of any results in the literature, though in principle they should follow easily from the high temperature expansion. In a forthcoming paper

we shall prove that Assumptions A.2 and A.3 follow from Assumption A.1 and some other very mild assumptions.

Theorem 1. *Suppose Assumption* A1 *holds. Let* $\theta(L)$ *be defined by*

$$\theta(L) = \sup_{f,\Lambda,\omega} \langle f; f \rangle_{\Lambda,\omega,\lambda} / \mathscr{D}_{\Lambda,\omega,\lambda}(f), \tag{1.20}$$

$$\mathscr{D}_{\Lambda,\omega,\lambda}(f) = \left\langle \sum_{x \in \Lambda} (\sigma_x f)^2 \right\rangle_{\Lambda,\omega,\lambda}. \tag{1.21}$$

Here $\langle \ \rangle_{\Lambda,\omega,\lambda} = E_{\Lambda,\omega,\lambda}$, $\langle u; v \rangle \equiv \langle uv \rangle - \langle u \rangle \langle v \rangle$ *and* Λ *is any generalized cube of size less than or equal to* L. *Then there is a constant* k *independent of* Λ, ω *and* λ *such that* $\theta(L) \leq k\varrho(1 - \varrho)$ *with* ϱ *denoting the density of the Gibbs state* $\mu_{\Lambda,\omega\lambda}$, *namely* $\varrho = \left\langle |\Lambda|^{-1} \sum_x \eta_x \right\rangle_{\Lambda,\omega,\lambda}$.

Theorem 2. *Suppose that Assumptions* A2 *and* A3 *hold and that Assumption* A1 *holds for all* λ. *Let*

$$w(L) = L^{-2} \sup_{f,\Lambda,\omega,N} \langle f; f \rangle_{\Lambda,\omega,N} / D_{\Lambda,\omega,N}(f), \tag{1.22}$$

$$D_{\Lambda,\omega,N}(f) = \left\langle \sum_{|x-y|=1} (T_{xy}f)^2 \right\rangle_{\Lambda,\omega,N}. \tag{1.23}$$

Here $\langle \ \rangle_{\Lambda,\omega,N} = E_{\Lambda,\omega,N}$. *Then there exists a constant* k *independent of* Λ, ω, *and* N *such that* $w(L) \leq k$.

The following Theorem 3 concerns the logarithmic Sobolev inequality for the Glauber dynamics (1.21). We first recall the definition of entropy. Let α and β be two probability measures. Then the entropy $S(\alpha/\beta)$ of α relative to β is defined by

$$S(\alpha/\beta) = \int \left[\log \left(\frac{d\alpha}{d\beta} \right) \right] d\alpha. \tag{1.24}$$

In the case we are interested in, (1.24) is well defined since both α and β will be discrete. In general one can define entropy by a variational principle.

Theorem 3. *Suppose Assumption* (A.1) *holds. Let* $u(L)$ *be defined by*

$$u(L) = \sup_{f,\omega,\Lambda} S(f\mu_{\Lambda,\omega,\lambda}/\mu_{\Lambda,\omega,\lambda}) / \mathscr{D}_{\Lambda,\omega,\lambda}(\sqrt{f}),$$

where the sup *is taken over all cubes with size less than or equal to* L *and all probability density (with respect to* $\mu_{\Lambda,\omega,\lambda}$) f. *Then there is a constant* k *independent of* Λ, ω *and* λ *such that* $u(L) \leq k$.

II. Proof of Theorem 1.

We shall prove Theorem 1 only for $d = 2$. The general case follows from similar arguments.

Step 1. Let Λ be the union of a generalized cube Ω_L with its translation $\tau_{(0,-L)}\Omega$, namely

$$\Lambda = \Omega_L \cup \tau_{(0,-L)}\Omega_L.$$

Define an order $x = (x^1, x^2) \le y = (y^1, y^2)$ if $x^2 < y^2$ or $x^2 = y^2$ and $x^1 \le y^1$. Let $x_0 = (0,0)$, $x_1 = (0,1)$ etc. and let $\eta_j = \eta_{x_j}$. For any j nonnegative define

$$f_j(\eta_j, \eta_{j+1}, \ldots) = E^\mu[f \mid \eta_j, \eta_{j+1} \ldots] = E[f \mid \mathscr{F}_j]. \tag{2.1}$$

Here $\mathscr{F}_j$ is the σ-algebra generated by $\eta_j, \eta_{j+1}, \ldots$. Then one has the identity

$$E[f; f] = \langle f; f \rangle = E\left\{ \sum_{j=0}^\infty \langle f_j; f_j \rangle_j + E[f; f \mid \mathscr{F}_0] \right\}. \tag{2.2}$$

Here $\langle f; f \rangle = \langle f^2 \rangle - \langle f \rangle^2$ and

$$\langle f_j; f_j \rangle_j = E[f_j^2 \mid \mathscr{F}_{j+1}] - E[f_j \mid \mathscr{F}_{j+1}]^2. \tag{2.3}$$

Note that the summation in (2.2) has only a finite number of terms.

Step 2. By definition f_j is the expectation of f with respect to the Gibbs measure with boundary condition ω and $\eta_j, \eta_{j+1} \ldots$. Let $\mu^{(j)}$ denote such a measure. Let $\nu^{(j)}$ denote the modified measure with boundary condition the same as $\mu^{(j)}$ except η_j is set to be 0. Let

$$d\mu^{(j)}/d\nu^{(j)} = h^{(j)}. \tag{2.4}$$

Then $(h = h^{(j)})$

$$f_j = \int fh \, d\nu^{(j)} = \int f \, d\nu^{(j)} + \langle f; h \rangle_{\nu^{(j)}} \equiv \widetilde{f}_j + \langle f; h \rangle_{\nu^{(j)}}. \tag{2.5}$$

Hence

$$\langle f_j; f_j \rangle_j \le 2\langle \widetilde{f}_j; \widetilde{f}_j \rangle_j + 2E[\langle f; h \rangle_{\nu^{(j)}}^2 \mid \mathscr{F}_{j+1}]. \tag{2.6}$$

Step 3. Given the σ-field $\mathscr{F}_{j+1}$, η_j is distributed according to some Bernoulli measure. Let p be the probability of having $\eta_j = 1$. Hence

$$\langle \widetilde{f}_j; \widetilde{f}_j \rangle_j = p(1-p)\big(\widetilde{f}_j(\eta_j = 1) - \widetilde{f}_j(\eta_j = 0)\big)^2. \tag{2.7}$$

It is straightforward to compute

$$(\widetilde{f}_j(\eta_j = 1) - \widetilde{f}_j(\eta_j = 0))^2 = \left(\int [f(\eta_j = 1) - f(\eta_j = 0)] \, d\nu_j \right)^2$$

$$\le \int (f(\eta_j = 1) - f(\eta_j = 0))^2 \, d\nu_j.$$

Clearly, there is a constant u so that $0 < u^{-1} < (p/\varrho) + (1-p)/(1-\varrho) < u < \infty$ with ϱ denoting the density defined in Theorem 1 and with u depending on the Hamiltonian but independent of λ. Together with (2.7) we have

$$E\langle \widetilde{f}_j; \widetilde{f}_j \rangle_j \le E\left\{ p(1-p) \int (\sigma_j f)^2 \, d\nu_j \right\}$$

$$\le u\varrho(1-\varrho)E\left\{ \int (\sigma_j f)^2 \, d\nu_j \right\}$$

$$\le \text{const. } \varrho(1-\varrho)E[(\sigma_j f)^2]. \tag{2.8}$$

Here we have used the fact that $h = h^{(j)}$ in (2.4) is bounded by a constant independent of λ.

Step 4. Let ℓ be a fixed large number and let $\widetilde{B}_\alpha$ be the cube (with j fixed and $\alpha = 0, 1, 2, \ldots$)

$$\widetilde{B}_\alpha = \{x \in \mathbb{Z}^d \mid \text{dist}(x, x_j) \le \ell^\alpha\}, \tag{2.10}$$

where x_j is the jth site according to our ordering in Step 1. Let $\widetilde{\mathscr{G}}_\alpha$ be the σ-algebra generated by $\{\eta_x \mid x \notin \widetilde{B}_\alpha\}$ and let $\mathscr{G}_\alpha = \widetilde{\mathscr{G}}_\alpha \cap \mathscr{F}_j$. Define $h_\alpha^{(j)}$ by

$$h_\alpha^{(j)} = E[h^{(j)} \mid \mathscr{G}_\alpha], \quad h_o^{(j)} = h^{(j)}. \tag{2.11}$$

Hence

$$\langle f; h^{(j)} \rangle_{\nu(j)}^2 = \left\langle f; \sum_{\alpha=0}^\infty \left(h_\alpha^{(j)} - h_{\alpha+1}^{(j)} \right) \right\rangle_{\nu(j)}^2$$

$$\le \text{const.} \sum_{\alpha=0}^\infty \langle f; h_\alpha^{(j)} - h_{\alpha+1}^{(j)} \rangle_{\nu(j)}^2 (\alpha+1)^2. \tag{2.12}$$

Note that by definition $E^{\nu(j)}\left[h_\alpha^{(j)} - h_{\alpha+1}^{(j)} \mid \mathscr{G}_{\alpha+1} \right] = 0$. Hence

$$\langle f; h_\alpha^{(j)} - h_{\alpha+1}^{(j)} \rangle_{\nu(j)}^2 = \{ E^{\nu(j)} E^{\nu(j)}[f; h_\alpha^{(j)} \mid \mathscr{G}_{\alpha+1}] \}^2$$

$$\le E^{\nu(j)} \{ E^{\nu(j)}[f; f \mid \mathscr{G}_{\alpha+1}] E^{\nu(j)}[h_\alpha^{(j)}; h_\alpha^{(j)} \mid \mathscr{G}_{\alpha+1}] \}. \tag{2.13}$$

Since $h^{(j)}$ is a local function at x_j, by Definition (2.11) and Assumption (A1) for any two configurations η and ζ

$$|h_\alpha^{(j)}(\eta) - h_\alpha^{(j)}(\zeta)| \le \text{const.} \, \ell^\alpha \exp[-C\ell^\alpha], \tag{2.14}$$

we have that

$$\langle f; h^{(j)} \rangle_{\nu(j)}^2 \le \sum_\alpha E^{\nu(j)} E^{\nu(j)}[f; f \mid \mathscr{G}_{\alpha+1}] \exp[-C\ell^\alpha](\alpha+1)^2 \ell^\alpha$$

$$\le \text{const.} \sum_\alpha E^{\mu(j)} E^\mu[f; f \mid \mathscr{G}_{\alpha+1}] \exp[-C\ell^\alpha](\alpha+1)^2 \ell^\alpha. \tag{2.15}$$

Here we have used the fact that $|d\nu^{(j)}/d\mu^{(j)}| \le \text{const.}$ with the constant independent of λ.

Step 5. Let $\alpha_1 > \alpha_0$ be chosen so that $\exp[-C\ell^{\alpha_1}] \le L^{-5d}$ and $\alpha_0^2 \ell^{\alpha_0} \exp[-C\ell^{\alpha_0}] \le \varepsilon$ for some ε small to be chosen later. Note that $\ell^{\alpha_1} \le \text{const.} \log L$. Divide the summation in (2.15) into three regions: $\alpha_1 > \alpha \ge \alpha_0$, $\alpha < \alpha_0$ and $\alpha \ge \alpha_1$. In the first region we use induction to have

$$E^\mu[f; f \mid \mathscr{G}_{\alpha+1}] \le \theta(\text{const.} \log L) E^\mu \left[\sum_{x \in \widetilde{B}_\alpha} (\sigma_x f)^2 \mid \mathscr{G}_{\alpha+1} \right].$$

For the second region, since α_0 is just a fixed constant, there is a $C(\alpha_0)$ such that

$$E^{\mu(j)} \sum_{\alpha < \alpha_0} E^\mu[f; f \mid \mathscr{G}_\alpha] \le C(\alpha_0)\varrho(1-\varrho) E^{\mu(j)} \left\{ \sum_{|x-x_j| \le \ell^{\alpha_0}} (\sigma_x f)^2 \right\}.$$

In the last region, one simply uses the trivial bound that $E^{\mu(j)}E^{\mu}[f; f \mid \mathscr{G}_{\alpha_1+1}] \leq E^{\mu(j)}[f; f]$. Hence we can bound $\langle f; h^{(j)} \rangle^2_{\nu(j)}$ by

$$
\begin{aligned}
\langle f; h^{(j)} \rangle^2_{\nu(j)} &\leq \theta(\text{const. log } L)\left\langle \sum_{\alpha_1>\alpha>\alpha_0}\left[\sum_{x\in\tilde{B}_\alpha}(\sigma_x f)^2\right]\right\rangle_{\mu(j)}(\alpha+1)^2\ell^\alpha \exp[-C\ell^\alpha] \\
&\quad + C(\alpha_0)\varrho(1-\varrho)\left\langle \sum_{|x-x_j|\leq\ell^{\alpha_0}}(\sigma_x f)^2\right\rangle_{\mu(j)} \\
&\quad + L^{-4d}\langle f; f\rangle_{\mu(j)}.
\end{aligned} \tag{2.17}
$$

Step 6. By induction,

$$
E[f; f \mid \mathscr{F}_0] \leq \theta(L)E\left[\sum_{x<x_0}(\sigma_x f)^2\right]. \tag{2.18}
$$

We now collect (2.2), (2.6), (2.8), (2.17) and choose ε small to have

$$
\begin{aligned}
\langle f; f\rangle_\Lambda &\leq \theta(L)\left\langle \sum_{x<x_0}(\sigma_x f)^2\right\rangle_\Lambda \\
&\quad + \left[\frac{1}{100}\theta(\text{const. log } L) + C(\alpha_0)\varrho(1-\varrho)\right]\left\langle \sum_x(\sigma_x f)^2\right\rangle_\Lambda \\
&\quad + L^{-2d}\langle f; f\rangle_\Lambda.
\end{aligned} \tag{2.19}
$$

Since we have a similar inequality if one reverses the order in Step 1 by reflection, we have the averaged inequality

$$
\begin{aligned}
\langle f; f\rangle_\Lambda &\leq \left[\frac{1}{2}\theta(L) + \frac{1}{100}\theta(\text{const. log } L) + C(\alpha_0)\varrho(1-\varrho)\right]\left\langle \sum_x(\sigma_x f)^2\right\rangle_\Lambda \\
&\quad + L^{-2d}\langle f; f\rangle_\Lambda.
\end{aligned} \tag{2.20}
$$

Step 7. Let $\tilde{\Lambda} = \Lambda \cup \tau_{(-L,0)}\Lambda$ be a cube of size $2L \times 2L$. Repeat the above procedure once more; we then conclude that

$$
\begin{aligned}
\theta(2L) &\leq \frac{2}{3}\theta(L) + \frac{1}{100}\theta(\text{const. log } L) + \text{constant } \varrho(1-\varrho) \\
&\leq \frac{3}{4}\theta(L) + \text{constant } \varrho(1-\varrho).
\end{aligned}
$$

The above inequality implies that $\theta(L)$ is bounded for all L and concludes Theorem 1. $\square$

III. Proof of Theorem 2, Part I

Our basic procedure for proving Theorem 2 is similar to that for proving Theorem 1. There are additional complications due to the conservation law and the slow decay of correlation functions (i.e., the $1/L^d$ term in Assumption A2). In this section, we shall bound $(f_j(\eta_j = 1) - f_j(\eta_j = 0))^2$ (see (2.1) for definition of f_j) by Dirichlet forms and covariances (cf. Steps 2 and 3 for Theorem 1). In the next section, we shall

bound the covariance by Dirichlet forms again (cf. Step 4 for Theorem 1) and thus conclude Theorem 2.

Let ν_N be a canonical Gibbs state on Λ with some fixed boundary condition and total number of particles N. For each $x \in \Lambda$ define

$$F_x = (1 - \eta_x) \exp\{-H(\sigma_x \eta) + H(\eta)\}. \tag{3.1}$$

Here H is the Hamiltonian for ν_N. Also define

$$F_N = (|\Lambda| - N)^{-1} \sum_{x \in \Lambda} F_x. \tag{3.2}$$

By particle-hole duality, we can assume without loss of generality that the density $\varrho_N = N/|\Lambda|$ is bounded by

$$\varrho_N < 2/3. \tag{3.3}$$

Assumption (3.3) will be enforced throughout the rest of this paper without further explanation.

Most results in this section hold trivially in the case of continuous dynamics, e.g. Ginzburg-Landau dynamics. Without going into the details of the Ginzburg-Landau dynamics, let us remark that the two basic operators σ_x (1.13) and T_{xy} (1.14) for the discrete dynamics will be replaced by

$$\sigma_x f = \partial f / \partial \eta_x,$$
$$T_{xy} f = \partial f / \partial \eta_x - \partial f / \partial \eta_y,$$

in the Ginzburg-Landau dynamics. Certainly in this case η_x is a continuous variable. As can be easily checked, the following Lemmas 3.1, 3.2, 3.4 and 3.5 are just simple consequences of chain rules for differentiation if it were the Ginzburg-Landau dynamics.

The reader should bear in mind that the discrete dynamics has to be treated carefully when the density $\varrho_N = N/|\Lambda|$ becomes very close to one or very close to zero. One certainly does not expect new phenomena occurred in this case; it nevertheless requires careful arguments to treat the discrete nature of our dynamics. More significantly, there are nontrivial differences between these two dynamics as we shall explain more carefully in Lemma 3.6.

Lemma 3.1. *With the above notation, for any function f*

$$E^{\nu_{N+1}}[f] - E^{\nu_N}[f]$$
$$= (N + 1)^{-1} \sum_x E^{\nu_{N+1}}[-(\sigma_x f)\eta_x] + [E^{\nu_N} F_N]^{-1} E^{\nu_N}[f; F_N]. \tag{3.4}$$

In particular, if f is a local function and (3.3) holds then

$$|[E^{\nu_{N+1}} - E^{\nu_N}][f]| \leq \text{const.}|\Lambda|^{-1}, \tag{3.5}$$

provided that Assumption A2 holds. Here σ_x is defined in (1.13).

Proof. By definition,

$$E^{\nu_{N+1}}[f] = (N+1)^{-1}E^{\nu_{N+1}}\left[\sum_x \eta_x f(\eta)\right]$$

$$= (N+1)^{-1}E^{\nu_{N+1}}\left[\sum_x \eta_x(f(\eta) - f(\sigma_x\eta))\right]$$

$$+ (N+1)^{-1}E^{\nu_{N+1}}\left[\sum_x \eta_x f(\sigma_x\eta)\right].$$

For each x fixed, change the variable $\eta \to \sigma_x\eta$. Thus

$$(N+1)^{-1}E^{\nu_{N+1}}\left[\sum_x \eta_x f(\sigma_x\eta)\right] = Z_{N+1}^{-1}Z_N(N+1)^{-1}E^{\nu_N}\left[\sum_x F_x(\eta)f(\eta)\right]$$

$$\equiv C_N E^{\nu_N}[F_N f]$$

$$= C_N E^{\nu_N}[F_N; f] + C_N E^{\nu_N}[F_N]E^{\nu_N}[f].$$

The constant C_N can be determined easily by putting $f = 1$:

$$1 = C_N E^{\nu_N}[F_N].$$

This proves (3.4). Using (3.4), Assumption A2 and (3.8) (to be proved in Lemma 3.3) we have (3.5) immediately.   □

Lemma 3.2 *Let $\nu_0 = \nu_{N,\omega_0}$ and $\nu_1 = \nu_{N,\omega_1}$ be two canonical Gibbs states with boundary condition ω_0 and ω_1 for which the only difference is that $(\omega_0)_z = 0$ while $(\omega_1)_z = 1$ for some $z \in \partial\Lambda$. Then for any function f*

$$E^{\nu_1}[f] - E^{\nu_0}[f] = E^{\nu_0}[f; h],$$

where $h = d\nu_1/d\nu_0$. In particular if f is a local function at x then

$$|[E^{\nu_1} - E^{\nu_0}][f]| \leq \text{const.}\,[L^{-d} + t(x-z)]\,\|f\|_\infty \tag{3.6}$$

provided that Assumption A2 holds. Furthermore if $f = \dfrac{1}{|\Lambda|}\sum_x f_x$ with f_x being a local function at x then

$$|[E^{\nu_1} - E^{\nu_0}][f]| \leq \text{const.}\,L^{-d}\|f\|_\infty, \tag{3.7}$$

where $\|f\|_\infty = \sup \|f_x\|_\infty$.

Proof. The identity before (3.6) is simply the definition of h. Inequality (3.6) follows from this identity and Assumption A2. Finally (3.7) follows from (3.6) and the fact t is summable.   □

Lemma 3.3. *For any $x \in \Lambda$ one has the bounds*

$$\text{const. } \varrho_N \leq E^{\nu_N}[\eta_x] \leq \text{const. } \varrho_N = \text{const. } N/|\Lambda|, \tag{3.8}$$
$$1 - \text{const. } (1 - \varrho_N) \leq E^{\nu_N}[\eta_x] \leq 1 - \text{const. } (1 - \varrho_N). \tag{3.9}$$

Furthermore, for any set $\{x_1, \ldots, x_k\} \subset \Lambda$ with $x_i \neq x_j$ whenever $i \neq j$ one has

$$(\text{const. } \varrho_N)^k \leq E^{\nu_N}[\eta_{x_1} \cdots \eta_{x_k}] \leq (\text{const. } \varrho_N)^k, \tag{3.10}$$
$$[1 - \text{const. } (1 - \varrho_N)]^k \leq E^{\nu_N}[\eta_{x_1} \cdots \eta_{x_k}] \leq [1 - \text{const. } (1 - \varrho_N)]^k. \tag{3.11}$$

Also for F_N defined in (3.2) one has

$$\text{const.} \leq E^{\nu_N}[F_N] \leq \text{const.} \tag{3.12}$$

with the constant depending only on the Hamiltonian.

Proof. First of all let us assume that (3.8) holds. Clearly by particle-hole duality one has that

$$\text{const. } (1 - \varrho_N) \leq E^{\nu_N}[1 - \eta_x] \leq \text{const. } (1 - \varrho_N).$$

It is elementary to check that this inequality is nothing but (3.9). Next we prove (3.10) for $k = 2$ assuming (3.8). The general case follows by induction.

$$\begin{aligned}
E^{\nu_N}[\eta_{x_1} \eta_{x_2}] &= E^{\nu_N}[\eta_{x_1}] E^{\nu_N}[\eta_{x_2} \mid \eta_{x_1} = 1] \\
&\leq \text{const. } \varrho N \cdot \text{const. } (N - 1)(|\Lambda| - 1)^{-1} \\
&\leq (\text{const. } \varrho_N)^2.
\end{aligned}$$

Similarly, one can prove (3.11) based on (3.9). So it remains to prove (3.8). For this purpose, it suffices to prove that for any two sites x and y,

$$E^{\nu_n}[\eta_x] \leq \text{const. } E^{\nu_N}[\eta_y].$$

This is because one can average (3.13) with respect to y to obtain (3.8). But this is a simple consequence of the fact that exchanging spins at x and y affects the Gibbs factor by at most some bounded factor. We have thus concluded Lemma 3.3. $\square$

Let z be a point in Λ and denote configurations in Λ by $\eta = (\eta_z, \xi)$. Let

$$H_0(\xi) = H(\eta_z = 0, \xi).$$

Denote by ν_0 the canonical Gibbs state with Hamiltonian H_0 and number of particles N. Let H' be the difference

$$H'(\eta_z, \xi) = H(\eta_z, \xi) - H_0(\xi),$$

and let g_z be defined by

$$g_z = \exp[-H'(\eta_z, \xi)]/E^{\nu_0}[\exp[-H'(\eta_z, \xi)] \mid \eta_z] = d\nu_N |_{\eta_z} / d\nu_0 |_{\eta_z}. \tag{3.13}$$

Lemma 3.4. *Recall the definition of* σ_x *and* T_{xy} *in Sect. 1. Then with the above notations,*

$$\{\sigma_z E^{\nu_N}[f \mid \eta_z]\}^2 \le 4E^{\nu_0}[f; g_z \mid \eta_z = 0]^2 + 4E^{\nu_0}[f; g_z \mid \eta_z = 1]^2$$
$$+ N^{-1} \sum_x E^{\nu_0}[\eta_x T_{zx} f \mid \eta_z = 0]^2$$
$$+ \text{const. } E^{\nu_0}[f; F_{N-1} \mid \eta_z = 1]^2. \tag{3.14}$$

Note that the left side of (3.14) *is independent of* η_z.

Remark. By definition $E[\cdot \mid \eta_z]^2 \equiv \{E[\cdot \mid \eta_z]\}^2$. We shall follow this convention for the rest of this paper.

Proof. By definition of g_z and ν_0,

$$E^{\nu_N}[f \mid \eta_z] = E^{\nu_0}[f g_z \mid \eta_z] = E^{\nu_0}[f; g_z \mid \eta_z] + E^{\nu_0}[f \mid \eta_z].$$

So by the Schwartz inequality

$$\{\sigma_z E^{\nu_N}[f \mid \eta_z]\}^2 = \{E^{\nu_0}[f; g_z \mid \eta_z = 0] - E^{\nu_0}[f; g_z \mid \eta_z = 1]$$
$$+ E^{\nu_0}[f \mid \eta_z = 0] - E^{\nu_0}[f \mid \eta_z = 1]\}^2$$
$$\le 4E^{\nu_0}[f; g_z \mid \eta_z = 0]^2 + 4E^{\nu_0}[f; g_z \mid \eta_z = 1]^2$$
$$+ 2\{\sigma_x E^{\nu_0}[f \mid \eta_z]\}^2.$$

Note that ν_0 depends on η_z only through the constraint $\bar{\xi} = N - \eta_z$. So we can apply Lemma 3.1 to $\sigma_z E^{\nu_0}[f \mid \eta_z]$. Therefore,

$$\{\sigma_z E^{\nu_0}[f \mid \eta_z]\}^2 \le 2N^{-1} \sum_x E^{\nu_0}[(T_{xz} f)\eta_x(1 - \eta_z) \mid \eta_z = 0]^2$$
$$+ 2E^{\nu_0}[F_{N-1} \mid \eta_z = 1]^{-2} E^{\nu_0}[f; F_{N-1} \mid \eta_z = 1]^2.$$

Note that we have changed σ_x to T_{xz} since the total number of particles is fixed. By (3.12), $E^{\nu_0}[F_{N-1} \mid \eta_z = 1] \ge \text{const.}$ This concludes Lemma 3.4.

Lemma 3.5. *With the same notation as in Lemma 3.4,*

$$E^{\nu_N}[E^{\nu_N}[f \mid \eta_z]; E^{\nu_N}[f \mid \eta_z]] \le 4E^{\nu_N}\{E^{\nu_0}[f; g_z \mid \eta_z]^2\}$$
$$+ \text{const.} \sum_x N^{-1} E^{\nu_0}[\eta_x(1 - \eta_z)T_{zx} f]^2$$
$$+ \text{const. } E^{\nu_N}\{\eta_z E^{\nu_0}[f; F_{N-1} \mid \eta_z]\}, \tag{3.15}$$

provided that (3.3) *holds.*

Proof. Let p denote $E^{\nu_N}\{\eta_z = 1\} = E^{\nu_N}[\eta_z]$. Since the marginal of ν_N on η_z is just a Bernoulli measure, we have

$$E^{\nu_N}[E^{\nu_N}[f \mid \eta_z]; \ E^{\nu_N}[f \mid \eta_z]]$$
$$= p(1 - p)\{E^{\nu_N}[f \mid \eta_z = 1] - E^{\nu_N}[f \mid \eta_z = 0]\}^2. \tag{3.16}$$

By Lemma 3.4,

$$
\begin{aligned}
E^{\nu_N}&[E^{\nu_N}[f \mid \eta_z];\; E^{\nu_N}[f \mid \eta_z]] \\
&= 4p(1-p)E^{\nu_N}E^{\nu_0}[f; g_z \mid \eta_z = 0]^2 \\
&\quad + 4p(1-p)E^{\nu_N}E^{\nu_0}[f; g_z \mid \eta_z = 1]^2 \\
&\quad + 4p(1-p)N^{-1}\sum_x E^{\nu_N}E^{\nu_0}[\eta_x T_{zx}f \mid \eta_z = 0]^2 \\
&\quad + \text{const.}\, p(1-p)E^{\nu_N}E^{\nu_0}[f; F_{N-1} \mid \eta_z = 1]^2 \\
&\leq 4E^{\nu_N}\{E^{\nu_0}[f; g_z \mid \eta_z]^2\} \\
&\quad + 4E^{\nu_N}\{(1 - \eta_z)N^{-1}\sum_x E^{\nu_0}[\eta_x T_{zx}f \mid \eta_z = 0]^2\} \\
&\quad + \text{const.}\, E^{\nu_N}\{\eta_z E^{\nu_0}[f; F_{N-1} \mid \eta_z]^2\}.
\end{aligned}
\tag{3.17}
$$

By definition of conditional expectation and (3.3)

$$
\begin{aligned}
E^{\nu_0}[\eta_x T_{zx}f \mid \eta_z = 0] &= E^{\nu_0}[\eta_x(1 - \eta_z)T_{zx}f]\{E^{\nu_0}[1 - \eta_z]\}^{-1} \\
&\leq \text{const.}\, E^{\nu_0}[\eta_x(1 - \eta_z)T_{zx}f].
\end{aligned}
$$

Hence the middle term of (3.17) is bounded by

$$
\text{const.} \sum_x N^{-1}[E^{\nu_0}\eta_x(1 - \eta_z)T_{zx}f]^2.
$$

This concludes Lemma 3.5. $\square$

Lemma 3.6. *Let ν_N be a canonical Gibbs state with total number of particles N in a cube Λ. Let γ be a path from x to z with $x, z \in \Lambda$ and with $\gamma_1 = x$ and $\gamma_{|\gamma|+1} = z$. For $i = 1, \ldots, |\gamma|$, let $b_i = (\gamma_i, \gamma_{i+1})$ be the bound connecting γ_i and γ_{i+1} and let*

$$
(T_{b_i}f)(\eta) = T_{\gamma_i, \gamma_{i+1}}f(\eta) = f(T_{\gamma_i, \gamma_{i+1}}\eta) - f(\eta).
\tag{3.18}
$$

Then there is a constant C depending only on the Hamiltonian such that

$$
\langle (f(T_{zx}\eta) - f(\eta))\eta_x(1 - \eta_z)\rangle^2_{\nu_N} \leq C|\gamma|\left\langle \sum_{i=1}^{|\gamma|}(T_{b_i}f)^2\right\rangle_{\nu_N},
\tag{3.19}
$$

provided that the density of particle $\varrho = N/|\Lambda|$ is strictly bounded away from one, say ϱ satisfies (3.3).

Remark. If one considers continuous spins with continuous dynamics (e.g. Ginzburg-Landau models), the bound (3.19) is just a simple consequence of the Schwartz inequality. The difficulty in Lemma 3.6 is completely due to the discrete nature of the dynamics. Should more than one species of particles per site be allowed (e.g., each site may have one black particle and one white particle), Lemma 3.6 still holds with only a slight modification of the arguments required.

Proof. Step 1. For any configuration η with $\eta_x = 1$ and $\eta_z = 0$ define a path $\Theta(\eta) = \{\theta_i(\eta)\}_{i=0}^{|\gamma|}$ connecting η to $T_{zx}\eta$ as follows.

Let $A(\eta)$ be the set of zeros for η, i.e. $A(\eta) = \{1 \leq i \leq |\gamma|+1 \mid \eta_i = 0\}$. Suppose $|A| = 1$. Then θ_i is defined by

$$
\theta_i\eta = T_{|\gamma|+1, |\gamma|+1-i}\eta, \quad i = 1, \ldots, |\gamma|, \quad \theta_0\eta = \eta.
$$

Here $T_{|\gamma|+1,|\gamma|+1-i}$ denote the operator exchanging particles at $\gamma_{|\gamma|+1}$ and $\gamma_{|\gamma|+1-i}$. Now assume that we have defined Θ for $|A| \leq k$ and η is a configuration with $|A(\eta)| = k + 1$. Let j be the smallest index in $A(\eta)$. Define

$$\theta_i \eta = T_{j,j-i}\eta, \quad \text{if} \quad i < j.$$

For $i \geq j$, let $\zeta = T_{1j}\eta$ and let δ denote the path from γ_j to $\gamma_{|\gamma|+1}$. Clearly, for the path δ the configuration ζ has only k zeros. Hence we can define θ_i for $i \geq j$ by using induction, namely

$$\theta_i(\eta) = \theta_{i-j}^{(\delta)}(T_{1j}\eta),$$

where the super index δ denote the dependence of θ on the path δ which has so far been omitted for the path γ. Note that $\theta_{|\gamma|}\eta = T_{xz}\eta$ and for each i there is a $u(i)$ so that

$$\theta_i(\eta) = T_{b_{u(i)}}\theta_{i-1}(\eta)$$

with T_b defined in (3.18). Furthermore, u is a bijection from $\{1, \ldots, |\gamma|\}$ to itself. Hence we have

$$f(T_{xz}\eta) - f(\eta) = \sum_{i=1}^{|\gamma|}[f(T_{b_{u(i)}}\theta_{i-1}(\eta)) - f(\theta_{i-1}(\eta))]$$

$$= \sum_{i=1}^{|\gamma|}[f(T_{b_i}\theta_{\alpha(i)-1}(\eta)) - f(\theta_{\alpha(i)-1})], \qquad (3.20)$$

where $\alpha = u^{-1}$. Denote the right side of (3.20) by

$$\sum_{i=1}^{|\gamma|} S_{b_i}f = f(T_{xz}\eta) - f(\eta). \qquad (3.21)$$

Let $\ell_i = \ell_i(\eta)$ denote the distance between $i + 1$ and the second zero after $i + 1$. By Schwartz' inequality and (3.21),

$$\langle(f(T^{zx}\eta) - f(\eta))\eta_x(1 - \eta_z)\rangle^2 \leq |\gamma| \sum_{i=1}^{|\gamma|}\langle(S_{b_i}f)^2\ell_i^{-1}\rangle\langle\eta_x(1 - \eta_z)\ell_i\rangle.$$

For each i fixed change the variable by $\theta_{\alpha-1}\eta = \xi$. Note that ξ differs from η by at most four sites. So the change of normalization and Boltzmann factor e^{-H} is bounded by some fixed constant. On the other hand, the mapping $\eta \to \theta_{\alpha-1}(\eta) = \xi$ is not one to one. For each ξ there may be more than one η with $\theta_{\alpha-1}(\eta) = \xi$. We now give an upper bound of the possible number of η with $\theta_{\alpha-1}(\eta) = \xi$. For simplicity we consider only $i = 1$. Let j be the position of second zero after 2. By construction, η and ξ agree after $j - 1$. The only source of confusion is that η has a zero between 2 and j and the position of this zero is arbitrary. Clearly the choice is bounded by $(j - 2)$. This proves the maximum number of η mapped into ξ is at most ℓ_i. Therefore

$$\langle(S_{b_i}f)\ell_i^{-1}\rangle \leq C\langle(T_{b_i}f)\rangle.$$

Step 2. To conclude Lemma 3.6, it remains to prove that

$$\langle\eta_x\ell_i\rangle \leq \text{const.}$$

Again for simplicity we assume that $i = 1$ and $\eta_1 = 1$. Clearly we only have to prove that

$$E[\eta_0\eta_1 \cdots (1 - \eta_j) \cdots \eta_s] \leq \text{const. } s^{-2}.$$

But this follows from (3.11). We have thus concluded Lemma 3.6. $\square$

For any two points $x = (x^1, x^2)$ and $y = (y^1, y^2)$ in Λ define the canonical path from x to y by first connecting x to (x^1, y^2) by a straight line and then connecting (x^1, y^2) to (y^1, y^2) again by a straight line. *From now on γ_{xy} will always denote the canonical path between x and y.* We can now combine Lemmas 3.5 and 3.6 to have

Lemma 3.7. *With the same notations and assumptions as Lemmas 3.5 and 3.6,*

$$E^{\nu_N}[E^{\nu_N}[f \mid \eta_z]; E^{\nu_N}[f \mid \eta_z]] \leq 4E^{\nu_N}\{E^{\nu_0}[f; g_z \mid \eta_z]^2\}$$

$$+ \text{const. } L|\Lambda|^{-1} \sum_x E^{\nu_N}\left[\sum_{i=1}^{|\gamma|}(T_{b_i}f)^2\right]$$

$$+ \text{const. } E^{\nu_N}\{\eta_z E^{\nu_0}[f; F_{N-1} \mid \eta_z]^2\}, \quad (3.22)$$

where $\gamma = \gamma_{zx}$.

Proof. Apply Lemma 3.6 with ν_N replaced by ν_0; we have (3.22) with the middle term on the right side replaced by E^{ν_0}. But switching ν_0 to ν_N costs at most some constant depending on the Hamiltonian. We have thus concluded (3.22). $\square$

Remark. Lemmas 3.3–3.7 are independent of Assumption A.2.

Corollary 3.8. *If, in addition, Assumption A.2 holds then*

$$E^{\nu_N}[E^{\nu_N}[f \mid \eta_z]; E^{\nu_N}[f \mid \eta_z]] \leq 4E^{\nu_N}\{E^{\nu_0}[f; g_z \mid \eta_z]^2\}$$

$$+ \text{const. } L|\Lambda|^{-1} \sum_x E^{\nu_N}\left[\sum_{i=1}^{|\gamma|}(T_{b_i}f)^2\right] + C(\varrho)|\Lambda|^{-1}E^{\nu_N}[f; f], \quad (3.23)$$

where $\gamma = \gamma_{zx}$ and $C(\varrho) \to 0$ and $\varrho \to 0$.

Proof. Use the Schwartz inequality and Assumption A.2 to bound the last term of (3.22). $\square$

IV. Proof of Theorem 2, Part II

Our goal in this section is to control the first and the last terms on the right side of (3.22). We then follow the same strategy as in the proof of Theorem 1 to conclude Theorem 2. The last term of (3.22) will be bounded in Lemma 4.4 while the first term will be bounded in Lemma 4.6. Lemmas 4.1–4.3 are preparations for Lemma 4.4. Lemma 4.4 (and its preparations Lemma 4.1–4.3) is the only place we need the assumption on three point functions, i.e., Assumption A3.

Let Λ be a cube of size L and $\Gamma \subset \Lambda$ be a subcube of size ℓ. Define the boundary $\partial \Gamma$ and $\partial^- \Gamma$ by

$$\partial \Gamma = \{x \in \Lambda \setminus \Gamma \mid |x - \Gamma| = 1\}, \quad (4.1)$$
$$\partial^- \Gamma = \{x \in \Gamma \mid |x - (\Lambda \setminus \Gamma)| = 1\}. \quad (4.2)$$

Let g be a function depending only on the configurations in Γ and let ν_N be a canonical Gibbs state on Λ. Define the function

$$g_\Gamma(\delta_0, \omega_0) = E^{\nu_N}[g \mid \omega = \omega_0, \ \delta = \delta_0], \tag{4.3}$$

where ω denotes the configuration on $\partial^- \Gamma$ and δ denotes the density in Γ.

Lemma 4.1. *Let g be either a local function at $x \in \Gamma$ or of the form*

$$g = |\Gamma|^{-1} \sum_{x \in \Gamma} g_x \tag{4.4}$$

with g_x a local function at x. Assume that g depends on the configurations inside Γ only. Then

$$E^{\nu_N}[g_\Gamma; g_\Gamma] \leq \text{const.} \, |\Gamma|^{-1}, \tag{4.5}$$

provided that $\ell < L^{1/20}$ and Assumptions A.2 and A.3 hold.

Proof. Let $\tilde{\nu}$ be the marginal of ν on Γ. By Corollary 5.6, $\tilde{\nu}$ has a spectral gap. Hence we only have to prove that

$$\|\sigma_x g_\Gamma\|_\infty \leq \text{const.} \, |\Gamma|^{-1}.$$

But this follows from (3.5) and (3.7). $\qquad\square$

Lemma 4.2. *Let $\nu_{N,\omega}$ be the canonical Gibbs state in a cube Λ with boundary condition ω and density $\varrho_N = N/|\Lambda| < 2/3$. Fix $z \in \partial\Lambda$, let $g = |\Lambda|^{-1} \sum_x g_x$, $\nu_{N,1} = \nu_{N,\omega}|_{\omega_z = 1}$ and $\nu_{N,0} = \nu_{N,\omega}|_{\omega_z = 0}$. Then*

$$|\{(E^{\nu_{N+1,1}} - E^{\nu_{N+1,0}}) - (E^{\nu_{N,1}} - E^{\nu_{N,0}})\}[g]| \leq C|\Lambda|^{-1}N^{-1}, \tag{4.6}$$

provided that Assumptions A.2 and A.3 hold. Also with the same assumption

$$|[E^{\nu_{N+1,\omega}} - 2E^{\nu_{N,\omega}} + E^{\nu_{N-1,\omega}}][g]| \leq C|\Lambda|^{-1}N^{-1}. \tag{4.7}$$

Proof. By Lemma 3.1,

$$
\begin{aligned}
&|\{E^{\nu_{N+1,1}} - E^{\nu_{N,1}}\} - \{E^{\nu_{N+1,0}} - E^{\nu_{N,0}}\}][g] \\
&= (N+1)^{-1} \sum_x \{E^{\nu_{N+1,1}}[-(\sigma_x g)\eta_x] - E^{\nu_{N+1,0}}[-(\sigma_x g)\eta_x]\} \\
&\quad + \{E^{\nu_{N,1}} F_N]^{-1} E^{\nu_{N,1}}[g; F_N] - [E^{\nu_{N,0}} F_N]^{-1} E^{\nu_{N,0}}[g; F_N]\} \\
&= \Omega_1 + \Omega_2.
\end{aligned}
$$

Step 1. We can rewrite Ω_2 as

$$
\begin{aligned}
\Omega_2 = -\{E^{\nu_{N,1}}[F_N]E^{\nu_{N,0}}[F_N]\}^{-1} \big[&\{E^{\nu_{N,1}}[F_N] - E^{\nu_{N,0}}[F_N]\}E^{\nu_{N,0}}[g; F_N] \\
&- E^{\nu_{N,0}}[F_N]\{E^{\nu_{N,1}}[g; F_N] - E^{\nu_{N,0}}[g; F_N]\} \big].
\end{aligned}
$$

Since $E^{\nu_N}[F_N]$ is bounded from above and below (3.12), Ω_2 is bounded by $C|\Lambda|^{-2}$ if one can prove that

$$
\begin{aligned}
\Omega_3 &= |E^{\nu_{N,1}}[F_N] - E^{\nu_{N,0}}[F_N]| \, |E^{\nu_{N,0}}[g; F_N]| \leq C|\Lambda|^{-2}, \\
\Omega_4 &= |E^{\nu_{N,1}}[g; F_N] - E^{\nu_{N,0}}[g; F_N]| \leq C|\Lambda|^{-2}.
\end{aligned}
$$

The bound on Ω_3 follows from (3.7) and Assumption A2. To bound Ω_4, let us assume without loss of generality that $E^{\nu_{N,0}}[F_N] = 0$. Denote the Radon-Nikodym derivative of $d\nu_{N,1}/d\nu_{N,0}$ by h_N. We can rewrite Ω_4 as

$$
\begin{aligned}
\Omega_4 &= E^{\nu_{N,1}}[g(\omega_z = 1)F_N] - E^{\nu_{N,1}}[g(\omega_z = 1)]E^{\nu_{N,1}}[F_N] \\
&\quad - E^{\nu_{N,0}}[g(\omega_z = 0)F_N] \\
&= E^{\nu_{N,0}}[g(\omega_z = 1)F_N h_N] - E^{\nu_{N,0}}[g(\omega_z = 1)F_N] \\
&\quad - E^{\nu_{N,0}}[g(\omega_z = 1)h_N]E^{\nu_{N,0}}[F_N h_N] \\
&\quad + E^{\nu_{N,0}}[\{g(\omega_z = 1) - g(\omega_z = 0)\}F_N].
\end{aligned}
$$

Denote $g(\omega_z = 1) = g_1$ and $g(\omega_z = 0) = g_0$. Then Ω_4 is just

$$
\begin{aligned}
\Omega_4 &= \{E^{\nu_{N,0}}[g_1 F_N; h_N] - E^{\nu_{N,0}}[g_1]E^{\nu_{N,0}}[F_N; h_N]\} \\
&\quad - E^{\nu_{N,0}}[g_1; h_N]E^{\nu_{N,0}}[F_N; h_N] + E^{\nu_{N,0}}[\{g_1 - g_0\}; F_N], \qquad (4.8)
\end{aligned}
$$

since $E^{\nu_{N,0}}[F_N] = 0$. The first term is equal to

$$
E^{\nu_{N,0}}[g_1; F_N; h_N].
$$

By Assumption A.3, it is bounded by $C|\Lambda|^{-2}$. By Assumption A.2 and definition of g the second and the third terms of (4.8) are bounded by $C|\Lambda|^{-2}$. We have thus proved that $\Omega_2 \le C|\Lambda|^{-2}$.

Step 2. It remains to bound Ω_1 by $C|\Lambda|^{-1}N^{-1}$. By similar arguments and notations as in Step 1,

$$
\begin{aligned}
\Omega_1 &= (N+1)^{-1}\sum_x E^{\nu_{N+1,0}}[(\sigma_x g_0)\eta_x] - E^{\nu_{N+1,1}}[(\sigma_x g_1)\eta_x] \\
&= (N+1)^{-1}\sum_x \{E^{\nu_{N+1,0}}[(\sigma_x g_1)\eta_x] - E^{\nu_{N+1,0}}[(\sigma_x g_1)\eta_x h_{N+1}]\} \\
&\quad + (N+1)^{-1}\sum_x E^{\nu_{N+1,0}}[\{(\sigma_x g_0) - (\sigma_x g_1)\}\eta_x] \\
&= -(N+1)^{-1}\sum_x E^{\nu_{N+1,0}}[(\sigma_x g_1)\eta_x; h_{N+1}] \\
&\quad + (N+1)^{-1}\sum_x E^{\nu_{N+1,0}}[\eta_x \sigma_x(g_0 - g_1)]. \qquad (4.9)
\end{aligned}
$$

By Assumption A.2 and the form of g, the two terms on the right side of (4.9) are bounded by const.$|\Lambda|^{-1}N^{-1}$. Hence $|\Omega_1| \le C|\Lambda|^{-1}N^{-1}$. Together with the bound on Ω_2 in Step 1, we have proved (4.7). The proof of (4.6) is similar and we omit it. $\square$

Lemma 4.3. *Let $\Gamma \subset \Lambda$ be a subcube with $\ell < L^{1/100}$. Denote by μ_λ the infinite volume Gibbs state with chemical potential λ. Let g be a function of the form $g = |\Gamma|^{-1}\sum_{x \in \Gamma} g_x$. Denote the density on Γ by ϱ_Γ. Let ν_ϱ be a canonical Gibbs state with number of particles $N = |\Lambda|\varrho$ and with $\varrho = p'(\lambda)$. Recall the definition of g_Γ in (4.3). Then for any $\varepsilon > 0$, $\gamma > 0$ and ℓ large enough*

$$
E^{\nu_\varrho}[g_\Gamma(\omega, \varrho_\Gamma) - k\varrho_\Gamma; g_\Gamma(\omega, \varrho_\Gamma) - k\varrho_\Gamma] \le \varepsilon|\Gamma|^{-1}, \qquad (4.10)
$$

provided that

$$
\mathrm{dist}(\Gamma, \Lambda^c) > L^{1/20} \qquad (4.11)
$$

and the density ϱ is bounded by

$$0 < \gamma < \varrho < 2/3. \tag{4.12}$$

Here k is a constant given by

$$k = |\Gamma| E^{\mu\lambda} [\widetilde{\sigma}_x g_\Gamma(\omega, \varrho_\Gamma)], \quad x \in \Gamma \setminus \partial^- \Gamma \tag{4.13}$$

with $\widetilde{\sigma}_x$ defined by

$$[\widetilde{\sigma}_x f](\eta) = [f(\sigma_x \eta) - f(\eta)](1 - \eta_x) - [f(\sigma_x \eta) - f(\eta)]\eta_x.$$

Note that the definition of k is independent of the choice of $x \in \Gamma \setminus \partial^- \Gamma$.

Proof. Step 1. Clearly by definition, $E[(\widetilde{\sigma}_x f)^2] = E[(\sigma_x f)^2]$. Note that we have the identity

$$\widetilde{\sigma}_x \varrho_\Gamma = |\Gamma|^{-1}.$$

By Corollary 5.6, the marginal of ν_N on Γ has a spectral gap. Hence

$$E^{\nu_N}[g_\Gamma(\omega, \varrho_\Gamma) - k\varrho_\Gamma; g_\Gamma(\omega, \varrho_\Gamma) - k\varrho_\Gamma]$$

$$\leq \text{const.} \sum_{x \in \Gamma} E^{\nu_N}[\{\widetilde{\sigma}_x(g_\Gamma(\omega, \varrho_\Gamma) - k\varrho_\Gamma)\}^2]$$

$$= \text{const.} \sum_{x \in \Gamma} E^{\nu_N}[\{\widetilde{\sigma}_x g_\Gamma(\omega, \varrho_\Gamma) - k/|\Gamma|\}^2].$$

Let us denote the contribution for $x \in \partial^- \Gamma$ by Ω_1 while $x \in \Gamma \setminus \partial^- \Gamma$ by Ω_2. By (3.7) Ω_1 is bounded by

$$\Omega_1 \leq \text{const.} |\Gamma|^{-2+(d-1)/d}, \tag{4.14}$$

where the factor $|\Gamma|^{(d-1)/d}$ comes from the summation of x over $\partial^- \Gamma$.

Step 2. We can write Ω_2 as

$$\Omega_2 = \text{const.} [|\Gamma| - |\partial^- \Gamma|] |\Gamma|^{-2} E^{\nu_N} [\widetilde{g}_\Gamma(\omega, \varrho_\Gamma) - k]^2,$$

where $\widetilde{g}_\Gamma$ is defined by

$$\widetilde{g}_\Gamma(\omega, \varrho) = |\Gamma| \widetilde{\sigma}_x g_\Gamma(\omega, \varrho), \quad x \in \Gamma \setminus \partial^- \Gamma.$$

Let $\alpha = E^{\nu_N}[\widetilde{g}_\Gamma(\omega, \varrho_\Gamma)]$. Then

$$\Omega_2 \leq \text{const.} |\Gamma|^{-1} E^{\nu_N}[\{\widetilde{g}_\Gamma(\omega, \varrho_\Gamma) - \alpha\}^2] + \text{const.} |\Gamma|^{-1}(k - \alpha)^2. \tag{4.15}$$

The first term is the variance of $\widetilde{g}_\Gamma$ and we can bound it again by using spectral gap as

$$\text{const.} |\Gamma|^{-1} E^{\nu_N} \left[\sum_{x \in \Gamma} \{\sigma_x \widetilde{g}_\Gamma(\omega, \varrho_\Gamma)\}^2 \right] \leq C(\varrho)|\Gamma|^{-2}. \tag{4.16}$$

Here we have used (4.6) in the last inequality. For the second term in (4.15), let A denote a subcube of size $L^{1/20}$ with the same center as $|\Gamma|$ and let $\tau_{\lambda,\omega}$ denote the Gibbs state on A with boundary condition ω. By the mixing Assumption A.1,

$$|E^{\tau_\lambda,\omega}[f] - E^{\tau_\lambda,\omega'}[f]| \leq \text{const.} L^{-1/20} \|f\|_\infty$$

if f depends only on configurations on Γ. Also, by the equivalence of ensembles Lemma 5.2,

$$|\alpha - k| = |E^{\nu_N}[\widetilde{g}_\Gamma(\omega, \varrho_\Gamma)] - E^{\mu_\lambda}[\widetilde{g}_\Gamma(\omega, \varrho_\Gamma)]| \leq \text{const.} L^{-1/20}. \tag{4.17}$$

Note that in order to apply Lemma 5.2 we have to change the expectation with respect to a Gibbs state in Λ by the infinite volume Gibbs state. The error we made is smaller than the right side of (4.17) by Assumption A.1. Lemma 4.3 follows from (4.14)–(4.17). $\quad\square$

Lemma 4.4. *Let Λ be a cube of size L or a cube of size L less a subcube of size smaller than $L^{1/100d}$. Let $\{g_x\}_{x\in\Lambda}$ be a family of translationally covariant functions in the sense that $\tau_{x-y}g_y = g_x$ unless x or/and y are close to the boundary. Assume that Assumption A1 holds for all λ. Then for any $0 < \ell < L$ there exists an $\varepsilon(\ell)$ with $\varepsilon(\ell) \to 0$ as $\ell \to \infty$ so that*

$$E^{\nu_N}\left[f; |\Lambda|^{-1}\sum_{x\in\Lambda} g_x\right]^2 \leq \text{const.}(\ell)|\Lambda|^{-1}D_\Lambda(f) + \varepsilon(\ell)|\Lambda|^{-1}\langle f; f\rangle_{\nu_N}, \qquad (4.18)$$

provided that L is large enough. Furthermore, $\text{const.}(\ell) \leq \exp(\text{const.}(\ell)$ for some constant. In the application we shall choose ℓ to be a fixed large constant independent of L.

Proof. First of all we can assume the density ϱ is strictly away from zero, i.e. $\frac{2}{3} > \varrho > \gamma > 0$. For otherwise,

$$E^{\nu_N}\left[f; |\Lambda|^{-1}\sum_{x\in\Lambda} g_x\right]^2 \leq E^{\nu_N}\left[|\Lambda|^{-1}\sum_{x\in\Lambda} g_x; |\Lambda|^{-1}\sum_{x\in\Lambda} g_x\right]E^{\nu_N}[f; f]$$

$$\leq C(\varrho)|\Lambda|^{-1}\|g\|_\infty^2 E^{\nu_N}[f; f].$$

Here we have used Assumption A.2. Since $C(\varrho) \to 0$ as $\varrho \to 0$, (4.18) holds if ϱ is sufficiently small. Hence we shall assume ϱ is bounded by

$$\tfrac{2}{3} > \varrho > \gamma > 0.$$

Step 1. Divide Λ into cubes of size ℓ with ℓ large but independent of L. Let α, β index such cubes and denote the typical cube by B_α. Let g_α denote

$$g_\alpha = \ell^{-d}\sum_{x\in B_\alpha} g_x.$$

Hence

$$E^{\nu_N}\left[f; |\Lambda|^{-1}\sum_{x\in\Lambda} g_x\right]^2 = E^{\nu_N}\left[f; \ell^d|\Lambda|^{-1}\sum_\alpha g_\alpha\right]^2$$

$$= \Omega_1 + \Omega_2,$$

$$\Omega_1 = 2E^{\nu_N}\left[f; \ell^d|\Lambda|^{-1}\sum_{|\alpha-\Lambda^c|\leq L^{1/4}} g_\alpha\right]^2,$$

$$\Omega_2 = 2E^{\nu_N}\left[f; \ell^d|\Lambda|^{-1}\sum_{|\alpha-\Lambda^c|>L^{1/4}} g_\alpha\right]^2. \qquad (4.19)$$

By the Schwartz inequality and Assumption A.2,

$$\Omega_1 \leq \frac{\varepsilon}{4}E^{\nu_N}[f; f]|\Lambda|^{-1}, \qquad (4.20)$$

if L is large enough. Hence Ω_1 is negligible for the purpose of (4.19). Let us redefine g_α by

$$g_\alpha = \ell^{-d} \sum_{|x - B^c_\alpha| > \ell^{1/4}} g_x \,.$$

The error term is again negligible by a similar argument. From now on, we shall not be very careful about the boundary terms as they are negligible in this lemma.

Step 2. Recall the definition of g_{B_α} in (4.3), i.e.

$$G_\alpha \equiv g_{B_\alpha} = E^{\nu_N}[g_\alpha \mid \omega_\alpha, \varrho_\alpha],$$

where ω_α denotes the configurations on $\partial^- B_\alpha$ and ϱ_α denotes the density at B_α. By the Schwartz inequality $\Omega_2 \leq \Omega_3 + \Omega_4$ with

$$\Omega_3 = 4 A v_\alpha \Omega_3^\alpha = 4 \ell^d |\Lambda|^{-1} \sum_\alpha E^{\nu_N}[f; g_\alpha - G_\alpha]^2,$$

$$\Omega_4 = 4 E^{\nu_N} \left[f; \ell^d |\Lambda|^{-1} \sum_\alpha G_\alpha \right]^2 .$$

We can bound Ω_3^α by

$$\Omega_3^\alpha = \{ E^{\nu_N} E^{\nu_N}[f; g_\alpha - G_\alpha \mid \omega_\alpha, \varrho_\alpha] \}^2$$
$$\leq E^{\nu_N} \{ E^{\nu_N}[f; f \mid \omega_\alpha, \varrho_\alpha] E^{\nu_N}[g_\alpha; g_\alpha \mid \omega_\alpha, \varrho_\alpha] \}.$$

By Assumption (A2), the second factor is bounded by $C|B_\alpha|^{-1}$. The first factor can be bounded by

$$\Omega_3^\alpha \leq \text{const.}(\ell) D_\alpha(f), \quad D_\alpha(f) = E^{\nu_N} \left[\sum_{b \in \alpha} (T_b f)^2 \right],$$

for some const.(ℓ) depending only on ℓ. Hence Ω_3 is bounded by

$$\Omega_3 \leq \text{const.}\, (\ell) |\Lambda|^{-1} D_\Lambda(f). \tag{4.21}$$

Step 3. Finally we have to bound Ω_4. Let $G_\alpha = G_\alpha^{(1)} + G_\alpha^{(2)}$, where

$$G_\alpha^{(1)} = \delta \varrho_\alpha$$

with δ a constant to be chosen later. Note that, thanks to the constraint $\sum \eta_x = $ constant, $\sum_\alpha G_\alpha^{(1)}$ is a constant. Hence we can replace G_α by $G_\alpha^{(2)}$ with arbitrary choice of δ. Hence

$$\Omega_4 \leq 4 \langle f; f \rangle \langle G^{(2)}; G^{(2)} \rangle, \tag{4.22}$$

where $G^{(2)} = \ell^d |\Lambda|^{-1} \sum_\alpha G_\alpha^{(2)}$. Note that by definition $\langle G^{(2)}; G^{(2)} \rangle = \langle G^{(2)}; G \rangle$ with $G = \ell^d |\Lambda|^{-1} \sum_\alpha G_\alpha$. Again by definition $(g = A v_\alpha g_\alpha)$,

$$\langle G^{(2)}; G^{(2)} \rangle = \langle G^{(2)}; G \rangle = \langle G^{(2)}; g \rangle. \tag{4.23}$$

For each α fixed, let

$$h_\alpha = E^{\nu_N}\left[|\Lambda|^{-1}\sum_{x\in\Lambda\setminus\Gamma} g_x \mid B_\alpha\right].$$

Then we can rewrite (4.22) as

$$\langle G^{(2)}; G^{(2)}\rangle = Av_\alpha\langle G_\alpha^{(2)}; h_\alpha\rangle$$
$$\leq Av_\alpha\langle G_\alpha^{(2)}; G_\alpha^{(2)}\rangle^{1/2}\langle h_\alpha; h_\alpha\rangle^{1/2}.$$

By Lemmas 3.1 and 3.2, $|\sigma_x h_\alpha| \leq \text{const.}|\Lambda|^{-1}$. Together with the spectral gap Corollary 5.3 we have

$$\langle h_\alpha; h_\alpha\rangle^{1/2} \leq \text{const.}|\Lambda|^{-1}\ell^{d/2}.$$

The first factor $\langle G_\alpha^{(2)}; G_\alpha^{(2)}\rangle$ can be bounded by

$$\langle G_\alpha^{(2)}; G_\alpha^{(2)}\rangle \leq \varepsilon\ell^{-d/2},$$

provided δ is chosen according to Lemma 4.3. Hence for ℓ large

$$\langle G^{(2)}; G^{(2)}\rangle \leq \frac{\varepsilon}{4}|\Lambda|^{-1}. \tag{4.24}$$

Together with (4.22) we have

$$\Omega_4 \leq \frac{\varepsilon}{4}|\Lambda|^{-1}\langle f; f\rangle.$$

Combining this bound with (4.21), we conclude Lemma 4.4. $\quad\square$

Corollary 4.5. *With assumptions and notation of Lemma 4.4. and Corollary 3.7,*

$$E^{\nu_N}[E^{\nu_N}[f \mid \eta_z]; E^{\nu_N}[f \mid \eta_z]] \leq |\Lambda|^{-1}LD_\Lambda(f) + \varepsilon|\Lambda|^{-1}E^{\nu_N}[f; f]$$
$$+ \text{const.}\, E^{\nu_N}\left\{L|\Lambda|^{-1}\sum_{y\in\Lambda}\sum_{i=1}^{|\gamma_{yz}|}(T_{b_i}f)^2\right\}$$
$$+ 4E^{\nu_N}\{E^{\nu_0}[f; g_z \mid \eta_z]^2\}. \tag{4.25}$$

Lemma 4.6. *Let h be a local function at $z \in \Lambda$ of size L. Then*

$$E^{\nu_N}[f; h]^2 \leq \text{const.}\, w(L^{1/100d})\sum_{b\in\Lambda}[1 + \text{dist}(b, z)]^{-d-1}E^{\nu_N}[(T_b f)^2]$$
$$+ \text{const.}\,|\Lambda|^{-1}L\sum_{y\in\Lambda}\sum_{x\in\Lambda}(1 + |y - x|)^{-d-1}\sum_{i=1}^{|\gamma_{yx}|}E^{\nu_N}[(T_{b_i}f)^2]$$
$$+ \varepsilon|\Lambda|^{-1}E^{\nu_N}[f; f]. \tag{4.26}$$

Here w is defined in (1.22).

Proof. Let $\widetilde{B}_\alpha$ denote a cube of size ℓ^α centered at z and let $B_\alpha = \widetilde{B}_\alpha \cap \Lambda$. Define (cf. 4.3)

$$h_\alpha(\omega_\alpha, \varrho_\alpha) = h_{B_\alpha} = E^{\nu_N}[h \mid \omega_\alpha, \varrho_\alpha].$$

We can bound the variance of h_α by Lemma 3.2, as

$$E^{\nu_N}[h_\alpha; h_\alpha] \leq \text{const.}\ell^{-d\alpha}.$$

Step 1. By the argument as in Step 2 of Lemma 4.4,

$$E^{\nu_N}[f; h]^2 \leq 2E^{\nu_N}[f; h - h_1]^2 + 2E^{\nu_N}[f; h_1]^2$$
$$\leq \text{const.}\, \ell^2 w(\ell) D_{B_1}(f) + 2E^{\nu_N}[f; h_1]^2. \qquad (4.27)$$

To bound the last term let $f^{(1)}$ denote

$$f^{(1)} = E^{\nu_N}[f \mid B_1].$$

Then by Assumption A.2,

$$\begin{aligned}
E^{\nu_N}[f; h_1]^2 &= E^{\nu_N}\left[f^{(1)}; h_1\right]^2 \\
&\leq E^{\nu_N}\left[f^{(1)}; f^{(1)}\right] E^{\nu_N}[h_1; h_1] \\
&\leq \text{const.}\, \ell^{-d} E^{\nu_N}\left[f^{(1)}; f^{(1)}\right]. \qquad (4.28)
\end{aligned}$$

By Corollary 5.6 ($\varrho = \varrho_N = N/|\Lambda|$),

$$E^{\nu_N}\left[f^{(1)}; f^{(1)}\right] \leq \text{const.}\, \varrho E^{\nu_N}\left[\sum_{x \in B_1} \left(\sigma_x f^{(1)}\right)^2\right].$$

We now apply Lemma 3.4 to $\sigma_x f^{(1)}$. Hence for $x \in \partial^- B_1$,

$$\begin{aligned}
E^{\nu_N}[(\sigma_x f^{(1)})^2] &\leq E^{\nu_N}\{E^{\nu_{N,x}}[f; g_x \mid \eta_x = 0, B_1]^2 + E^{\nu_{N,x}}[f; g_x \mid \eta_x = 1, B_1]^2 \\
&\quad + (N - N_{B_1} - 1)^{-1} \sum_{y \in \Lambda \backslash B_1} E^{\nu_{N,x}}[\eta_y T_{xy} f \mid \eta_x = 0, B_1]^2 \\
&\quad + \text{const.}\, E^{\nu_{N,x}}[f; F_{N-N_{B_1}-1} \mid \eta_x = 1]^2\}.
\end{aligned}$$

Here N_{B_1} denote the number of particles in B_1 and $\nu_{N,x}$ denotes the measure with the Hamiltonian having $\eta_x = 0$ (cf. definition of ν_0 in Lemma 3.4). Similarly, we have the same bound for $x \in B_1 - \partial^- B_1$ except the first two terms disappeared.

We can now proceed as in Lemmas 3.5–3.7 and Corollary 4.5. Note that the factor p in (3.16) which is essential for (3.17) to hold is supplied here by the factor ϱ. To summarize, we have

$$\begin{aligned}
\ell^{-d} E^{\nu_N}\left[f^{(1)}; f^{(1)}\right] &\leq \Omega_1 + \varepsilon |\Lambda|^{-1} E^{\nu_N}[f; f] \\
&\quad + \text{const.}\, |\Lambda|^{-1} L E^{\nu_N}\left\{\sum_{y \in B_1} \sum_{x \in \Lambda \backslash B_1} \sum_{i=1}^{|\gamma_{yx}|} (T_{b_i} f)^2\right\}, \qquad (4.29)
\end{aligned}$$

where Ω_1 is given by

$$\Omega_1 = \text{const.}\, \ell^{-d} E^{\nu_N}\left\{\sum_{x \in \partial^- B_1} E^{\nu_{N,x}}[f; g_x \mid B_1]^2\right\}. \qquad (4.30)$$

Together with (4.27) we have

$$E^{\nu_N}[f; h] \leq \Omega_1 + \varepsilon |\Lambda|^{-1} E^{\nu_N}[f; f] + \text{const.}\, \ell^2 w(\ell) D_{B_1}(f) + U_1, \qquad (4.31)$$

where U_1 is defined by

$$U_1 = \text{const.}\, |\Lambda|^{-1} L \ell^{-d} E^{\nu_N} \left\{ \sum_{y \in B_1} \sum_{x \in \Lambda \setminus B_1} \sum_{i=1}^{|\gamma_{yx}|} (T_{b_i} f)^2 \right\}. \qquad (4.32)$$

Step 2. Repeat Step 1 with Λ replaced by $\Lambda \setminus B_1$ and B_1 replaced by B_2. We can bound Ω_1 by

$$\Omega_1 \leq \Omega_2 + \text{const.}\, \varepsilon \ell^{-1} |\Lambda \setminus B_1|^{-1} E^{\nu_N} E^{\nu_N}[f; f \mid B_1]$$

$$+ \text{const.}\, \ell^{-2d-1} |\Lambda \setminus B_1|^{-1} L E^{\nu_N} \left\{ \sum_{y \in B_2} \sum_{x \in \Lambda \setminus B_2} \sum_{i=1}^{|\gamma_{yx}|} (T_{b_i} f)^2 \right\}$$

$$+ \text{const.}\, \ell^{2 \cdot 2 - 1} w(\ell^2) D_{B_2}(f), \qquad (4.33)$$

where Ω_2 is given by

$$\Omega_2 = \text{const.}\, \ell^{-2d-1} E^{\nu_N} \sum_{x \in \partial^- B_2} E^{\nu_N, x}[f; h_x \mid B_2]^2. \qquad (4.34)$$

One can replace the covariance in the middle term $E^{\nu_N} E^{\nu_N}[f; f \mid B_1]$ by $E^{\nu_N}[f; f]$ to have an upper bound. Also choose ℓ so large that $2\text{const.} \leq \ell^{1/2}$. Hence we can now rewrite (4.33) as

$$\Omega_1 \leq \Omega_2 + \varepsilon \ell^{-1/2} |\Lambda|^{-1} E^{\nu_N}[f; f]$$

$$+ \ell^{-2d-1/2} |\Lambda|^{-1} L E^{\nu_N} \left\{ \sum_{y \in B_2} \sum_{x \in \Lambda \setminus B_2} \sum_{i=1}^{|\gamma_{yx}|} (T_{b_i} f)^2 \right\}$$

$$+ \ell^{2 \cdot 2 - 1/2} w(\ell^2) D_{B_2}(f), \qquad (4.35)$$

where Ω_2 satisfies

$$\Omega_2 \leq \ell^{-2d-1/2} E^{\nu_N} \sum_{x \in \partial^- B_2} E^{\nu_N, x}[f; h_x \mid B_2]^2. \qquad (4.36)$$

We now repeat the same argument until we reach α_0 with $\ell^{\alpha_0} = L^{(100d)^{-1}}$. Hence we have

$$E^{\nu_N}[f; g] \leq \Omega_{\alpha_0} + 2\varepsilon |\Lambda|^{-1} E^{\nu_N}[f; f]$$

$$+ L|\Lambda|^{-1} \sum_{\alpha=1}^{\alpha_0} \ell^{-d\alpha - (1/2)(1+2+\cdots+\alpha-1)} \sum_{y \in B_\alpha} \sum_{x \in \Lambda \setminus B_\alpha} E^{\nu_N} \left\{ \sum_{i=1}^{|\gamma_{yx}|} (T_{b_i} f)^2 \right\}$$

$$+ w(\ell^{\alpha_0}) \left[\sum_{\alpha=1}^{\alpha_0} \ell^{2\alpha - (1/2)(1+2+\cdots+\alpha-1)} D_{B_\alpha}(f) \right] \qquad (4.37)$$

with

$$\Omega_{\alpha_0} \leq \ell^{-\alpha_0 d - (1/2)(1+2+\cdots+\alpha_0-1)} E^{\nu_N} \sum_{x \in \partial^- B_{\alpha_0}} E^{\nu_N, x}[f; g_x \mid B_{\alpha_0}]^2$$

$$\leq \text{const.}\, \ell^{-\alpha_0 d - (\alpha_0-1)\alpha_0/4} \cdot \ell^{\alpha_0(d-1)} E^{\nu_N}[f; f]. \tag{4.38}$$

Here we have used the Schwartz inequality in the last inequality. Since $\ell^{\alpha_0} = L^{(100d)^{-1}}$, for L large the last factor $\ell^{-2\alpha_0 - (\alpha_0-1)\alpha_0/4} < L^{-2d}$. So we can absorb Ω_{α_0} into the second term on the right side of (4.37). For the third term note that for each $y \in B_\alpha \setminus B_{\alpha-1}$, the numerical factor is

$$\sum_{\alpha=a}^{\alpha_0} \ell^{-d\alpha - (1/2)(1+2+\cdots+\alpha-1)} \leq \text{const.}\,(\ell)[1 + \ell^{(a-1)}]^{-d-1}. \tag{4.39}$$

So we can bound the second term as

$$\text{const.}\,(\ell)L|A|^{-1} \sum_{y \in A} \sum_{x \in A} (1 + |y - x|)^{-d-1} \sum_{i=1}^{|\gamma_{yx}|} E^{\nu_N}[(T_{b_i}f)^2]. \tag{4.40}$$

By similar arguments, the last term is bounded by

$$\text{const.}\,(\ell)w(L^{1/100d}) \sum_{b \in A} E^{\nu_N}[(1 + \text{dist}(b, z))^{-d-1}(T_b f)^2]. \tag{4.41}$$

Combining (4.37)–(4.41), we have concluded Lemma 4.6. $\quad\square$

We can now use Lemma 4.6 in Corollary 4.5 to obtain the following corollary.

Corollary 4.7.

$$E^{\nu_N}\{E^{\nu_N}[f \mid \eta_z]; E^{\nu_N}[f \mid \eta_z]\}$$

$$\leq \text{const.}\,\left\{ |A|^{-1} L D_A(f) + \varepsilon |A|^{-1} E^{\nu_N}[f; f] \right.$$

$$+ E^{\nu_N}\left\{ |A|^{-1} L \sum_{y \in A} \sum_{x \in A} (1 + |y - x|)^{-d-1} \sum_{i=1}^{|\gamma_{yx}|} (T_{b_i}f)^2 \right\}$$

$$\left. + w(L^{1/2}) \sum_{b \in A} [1 + |b - z|]^{-d-1} E^{\nu_N}[(T_b f)^2] \right\}.$$

Proof of Theorem 2. Step 1. Assume that (1.22) holds for cubes of size not bigger than L. Let $A = \Omega_L \cup \tau_{(0,-L)}\Omega_L = \Omega^1 + \Omega^2$ (cf. Step 1 in the proof of Theorem 1). We shall prove that

$$E^A[f; f] \leq \left(\tfrac{2}{3}w(L) + \text{const.}\right) L^2 D_{\Omega_L}(f). \tag{4.42}$$

As in the Step 7 of the proof of Theorem 1, we repeat the argument for $\widetilde{A} = A \cup \tau_{(-L,0)}A$. Hence

$$w(2L) \leq \tfrac{2}{3}w(L) + \text{const.}$$

Therefore $w(L)$ is uniformly bounded. We now prove (4.42).

Step 2. Recall (2.1)–(2.3). We now apply Corollary 4.7 to bound $E[f_k; f_k \mid \mathscr{F}_{k+1}]$ and the sum over k to have

$$\langle f; f \rangle_\Lambda \leq E[f; f \mid \mathscr{F}_1] + [\varepsilon w(L) + \text{const.}]L^2 D_\Lambda(f) + \varepsilon \langle f; f \rangle_\Lambda$$
$$\leq w(L)L^2 D_{\Omega^1}(f) + [\varepsilon w(L) + \text{const.}]L^2 D_\Lambda(f) + \varepsilon \langle f; f \rangle_\Lambda.$$

Switch the role of Ω^1 and Ω^2 and average,

$$(1 - 2\varepsilon)\langle f; f \rangle_\Lambda \leq \left[\left(\tfrac{1}{2} + \varepsilon\right)w(L) + \text{const.}\right]L^2 D_\Lambda(f).$$

This proves (4.42) and concludes Theorem 2. $\quad\square$

V. Equivalence of Ensemble

In this section, we shall prove a strong version of equivalence of ensemble based on the mixing Assumption A.2. The equivalence of ensemble is an old subject and has been studied extensively in the literature, see e.g. [R]. But most classical results are too weak for our purpose. We are able to obtain stronger estimates because of the mixing Assumption A.2.

Lemma 5.1 *Let $\lambda \in \mathbb{R}$ be a chemical potential. Then the pressure $p_{\Lambda,\omega}(\lambda)$ in a cube Λ of size L satisfies*

$$|p_{\Lambda,\omega}(\lambda) - p(\lambda)| \leq C/L \tag{5.1}$$

with C independent of L or λ or the boundary condition ω. Similarly, let ϱ be the density of particles defined by $\varrho = N/|\Lambda|$, then with the same constant C the free energy $f_{\Lambda,\omega}$ satisfies

$$f_{\Lambda,\omega}(\varrho) \geq f(\varrho) + C/L. \tag{5.2}$$

Proof. The proof for the pressure is obvious and we omit it. The free energy bound is also obvious since

$$f_{\Lambda,\omega}(\varrho) \geq \sup_\lambda(\lambda\varrho - p_{\Lambda,\omega}(\lambda))$$
$$\geq \sup_\lambda(\lambda\varrho - p(\lambda)) + C/L = f(\varrho) + C/L. \quad\square$$

Lemma 5.2. *Let Λ be a cube in $\mathbb{Z}^d$ and $\Gamma \subset \Lambda$ be a subcube. Denote the configurations on Λ by $\theta = (\eta, \zeta)$ with η denoting configurations on $\Lambda \setminus \Gamma$ and ζ configurations on Γ. Let $U_\lambda(\zeta)$ $(V_\varrho(\zeta))$ denote the marginal density of ζ with respect to the (canonical) Gibbs state with the boundary condition ω and chemical potential λ (density $\varrho = p'(\lambda)$). Then there is a constant C independent of ϱ, λ or ω so that*

$$|V_\varrho(\zeta) - U_\lambda(\zeta)| \leq CL^{-1/3}U_\lambda(\zeta), \tag{5.3}$$

provided that Assumption (A2) holds and

$$|\Gamma| < L^{1/20}, \quad L^{-1/8} < \varrho < 9/10. \tag{5.4}$$

Proof. Step 1. By definition $(\bar\eta = \#$ of η particles$)$

$$V_{n+1}(\zeta) = Z_{n+1}^{-1} \sum_\eta \exp[-H(\eta, \zeta)]\delta(\bar\eta + \bar\zeta = n+1)$$

$$= Z_{n+1}^{-1} \sum_\eta (n+1-\bar\zeta)^{-1}$$

$$\times \sum_{x \in \Lambda \setminus \Gamma} \eta_x \exp[-H(\eta, \zeta)]\delta(\bar\eta + \bar\zeta = n+1). \qquad (5.5)$$

Here we have abused the notation and use V_{n+1} for $V_{n+1/|\Lambda|}$. For each x fixed, let I_x be defined by

$$I_x(\theta) = (1 - \theta_x)\exp[-H(\sigma_x\theta) + H(\theta)], \qquad (5.6)$$

where σ_x is defined in (1.15). For each x fixed, change the variable by $\eta = \sigma_x\xi$. Thus

$$\sum_\eta \eta_x \exp[-H(\widetilde\theta)]\delta(\bar\eta + \bar\zeta = n+1) = \sum_\xi I_x(\theta)\exp[-H(\theta)]\delta(\bar\xi + \bar\zeta = n), \quad (5.7)$$

where $\widetilde\theta = (\eta, \zeta)$ on the left side of (5.7) and $\theta = (\xi, \zeta)$ on the right side. Denote by I

$$I(\theta) = (n+1-\bar\zeta)^{-1} \sum_{x \in \Lambda \setminus \Gamma} I_x(\theta). \qquad (5.8)$$

We can summarize Step 1 by

$$V_{n+1}(\zeta) = V_n(\zeta)Z_n Z_{n+1}^{-1} E^{\nu_n}[I \mid \zeta]. \qquad (5.9)$$

Step 2. Rewrite (5.9) as

$$V_{n+1}(\zeta) - V_n(\zeta) = V_n[Z_n Z_{n+1}^{-1} E^{\nu_n}[I \mid \zeta] - 1]. \qquad (5.10)$$

Also let $\Gamma = \emptyset$ we have

$$Z_{n+1} = Z_n E^{\nu_n}[G_{n+1}]. \qquad (5.11)$$

Here G_{n+1} is defined by

$$G_{n+1}(\theta) = (n+1)^{-1} \sum_{x \in \Lambda} I_x.$$

Using (5.11) in (5.10) we have that

$$V_{n+1}(\zeta) - V_n(\zeta) = V_n[E^{\nu_n}G_{n+1}]^{-1}\{E^{\nu_n}[I \mid \zeta] - E^{\nu_n}[G_{n+1}]\}. \qquad (5.12)$$

Step 3. By definition of I_x, there is a constant k such that

$$k^{-1}(1 - \theta_x) \leq I_x \leq k(1 - \theta_x). \qquad (5.13)$$

Together with (5.4), we can bound $E^{\nu_n}[G_{n+1}]$ by

$$k|\Lambda|n^{-1} \geq E^{\nu_n}[G_{n+1}] \geq k^{-1}[(|\Lambda| - n)(n+1)^{-1}] \geq \text{const.}\, k^{-1}|\Lambda|n^{-1}. \qquad (5.14)$$

We are now ready to bound $E^{\nu_n}[I|\zeta] - E^{\nu_n}[G_{n+1}]$ in (5.12). Let us first decompose the summation into $x \in \Gamma$ and $x \notin \Gamma$, namely,

$$
\begin{aligned}
E^{\nu_n}[I|\zeta_0] - E^{\nu_n}[G] = {} & -(n+1)^{-1} \sum_{x \in \Gamma} E^{\nu_n}[I_x] \\
& + (n+1-\bar\zeta_0)^{-1} \sum_{x \notin \Gamma} \{ E^{\nu_n}[I_x \mid \zeta_0] - E^{\nu_n}[I_x] \} \\
& + [(n+1-\bar\zeta_0)^{-1} - (n+1)^{-1}] \sum_{x \notin \Gamma} E^{\nu_n}[I_x]. \quad (5.15)
\end{aligned}
$$

By (5.13), (5.14) and (5.4) the first and the third term together are bounded by

$$
\text{const.} \{ |\Gamma| n^{-1} + |\Gamma| |\Lambda| n^{-2} \} \le \text{const.} |\Gamma| |\Lambda| n^{-2}. \quad (5.16)
$$

Combining (5.12)–(5.16) and (5.4) one has the following bound:

$$
|V_{n+1}(\zeta_0) - V_n(\zeta_0)| \le \text{const.} \, L^{1/8} V_n(\zeta_0) \{ |\Gamma| n^{-1} + |\gamma| \}, \quad (5.17)
$$

$$
\gamma = |\Lambda|^{-1} \sum_{x \notin \Gamma} \{ E^{\nu_n}[I_x \mid \zeta_0] - E^{\nu_n}[I_x] \}. \quad (5.18)
$$

Certainly $|\gamma|$ is bounded by

$$
\delta = \sup_{\zeta_1, \zeta_2} \left| E^{\nu_n}[|\Lambda|^{-1} \sum_{x \notin \Gamma} I_x \mid \zeta_1] - E^{\nu_n}[|\Lambda|^{-1} \sum_{x \notin \Gamma} I_x \mid \zeta_2] \right|,
$$

where the sup is taken over any two configurations on Γ. Note that $E^{\nu_n}[\cdot \mid \zeta_1]$ and $E^{\nu_n}[\cdot \mid \zeta_2]$ are canonical Gibbs measures on $\Lambda \setminus \Gamma$. We can now apply Lemma 3.2 to have

$$
|\gamma| \le \delta \le \text{const.} |\Gamma| |\Lambda|^{-1}.
$$

Therefore, one has the bound

$$
|V_{n+1}(\zeta) - V_n(\zeta)| \le \text{const.} \, L^{1/8} V_n(\zeta) |\Gamma| n^{-1}. \quad (5.19)
$$

By induction, for any n and m with $9|\Lambda|/10 \ge m \ge n \ge |\Lambda| L^{-1/8}$ one has

$$
|V_m - V_n| \le \text{const.} \, |m - n| L^{1/8} |\Gamma| n^{-1} V_n. \quad (5.20)
$$

The restriction $m \ge n$ can be removed provided that $|m - n| L^{1/8} |\Gamma| \ll n$. Hence (5.20) holds if

$$
|m - n| < |\Lambda| L^{-2/5}, \quad 9|\Lambda|/10 \ge n \ge |\Lambda| L^{-1/8}. \quad (5.21)
$$

Step 4. By definition,
$$
U_\lambda(\zeta) - V_\varrho(\zeta) = (\Omega_1 + \Omega_2) V_\varrho,
$$

where Ω_1 and Ω_2 are defined by ($\varrho_N \equiv N/|\Lambda|$)

$$
\Omega_1 = \sum_{|\varrho_N - \varrho| < L^{-2/5}} Z_N^c Z_\lambda^{-1} (V_{\varrho_N} - V_\varrho) V_\varrho^{-1} e^{\lambda N},
$$

$$
\Omega_2 = \sum_{|\varrho_N - \varrho| > L^{-2/5}} Z_N^c Z_\lambda^{-1} (V_{\varrho_N} - V_\varrho) V_\varrho^{-1} e^{\lambda N}.
$$

By (5.20) and (5.14) Ω_1 is bounded by

$$|\Omega_1| \leq \text{const.} |\Gamma| L^{1/4} \sum_{|\varrho_N - \varrho| < L^{-2/5}} Z_N Z_\lambda^{-1} e^{\lambda N} |\varrho_N - \varrho|$$

$$\leq \text{const.} |\Gamma| L^{-3/20} \leq \text{const.} L^{-1/10}. \tag{5.22}$$

To bound Ω_2, note that by (5.1) and (5.2),

$$Z_N Z_\lambda^{-1} e^{\lambda N} = \exp[|\Lambda| \{\lambda \varrho_N - f_\Lambda(\varrho_N) - p_\Lambda(\lambda)\}]$$

$$\leq \exp[|\Lambda| \{\lambda \varrho_N - f(\varrho_N) - p(\lambda) + CL^{-1}\}].$$

Since f is strictly convex in the sense that $f''(x) \geq \varepsilon > 0$, one has

$$\lambda x - f(x) - p(\lambda) \leq -\varepsilon(x - \varrho)^2.$$

So for ϱ_N with $|\varrho_N - \varrho| > L^{-2/5}$ one has

$$Z_N Z_\lambda^{-1} e^{\lambda N} < \exp\{-\text{const.} L^{d-1}\}. \tag{5.23}$$

Since V_{ϱ_N} is a probability density on the configurations on Γ, it follows from (3.10) and (3.11) that

$$|V_{\varrho_N}(\varsigma) V_\varrho^{-1}(\varsigma)| \leq |V_\varrho^{-1}(\varsigma)| \leq \text{const.} [\varrho^{-1}]^{|\Gamma|}.$$

Together with (5.23) we have $\Omega_2 < |\Lambda|^{-1}$. Lemma 5.2 follows from this bound and (5.22). □

Corollary 5.3. *Assume the notations and assumptions of Lemma 5.2. Suppose μ_λ has a spectral gap δ_λ. Then V_ϱ has a gap at least $\delta_\lambda/2$ (with respect to Glauber dynamics).*

Proof. By definition

$$\langle f; f \rangle_{V_\varrho} = \sum_\eta \sum_\varsigma [f(\eta) - f(\varsigma)]^2 V_\varrho(\eta) V_\varrho(\varsigma)$$

$$\leq 1.5 \sum_\eta \sum_\varsigma [f(\eta) - f(\varsigma)]^2 U_\lambda(\eta) U_\lambda(\varsigma)$$

$$\leq 1.5 \delta_\lambda \sum_\eta [\sigma_x f]^2(\eta) U_\lambda(\eta)$$

$$\leq 2\delta_\lambda \sum_\eta [\sigma_x f]^2(\eta) V_\varrho(\eta).$$

This concludes Corollary 5.3. □

Corollary 5.3 gives a sufficient condition for which the marginal density V has a spectral gap. But the result fails when the density ϱ becomes very close to zero. On the other hand, the spectral gap in the extremely low density case should be obvious as it corresponds to Gibbs measure with very high magnetic field. The following Lemmas 5.4 and 5.5 provide a simple sufficient condition to close this gap.

Lemma 5.4. *Suppose W is a probability density on Γ satisfying ($\bar{\eta}$ = # of η particles)*

$$\eta_x W(\eta) \le \alpha W(\sigma_x \eta), \quad \alpha |\Gamma| < 1/4, \tag{5.24}$$

for some positive constant α. Then W has a spectral gap at least $16/\alpha$ in the sense that

$$\langle f; f \rangle_W \le \alpha/16 \mathscr{D}_\Gamma(f). \tag{5.25}$$

Here $\mathscr{D}_\Gamma(f)$ is defined by (1.21) with the underlying measure Γ.

Proof. By definition and Schwartz inequality,

$$\begin{aligned}
\langle f; f \rangle_W &= \sum_\eta \sum_\zeta [f(\eta) - f(\zeta)]^2 W(\eta) W(\zeta) \\
&\le 4 \sum_\eta \sum_\zeta [f(\eta) - f(0)]^2 W(\eta) W(\zeta) \\
&\le 4 \sum_{\eta \ne 0} [f(\eta) - f(0)]^2 W(\eta).
\end{aligned} \tag{5.26}$$

The last term can be bounded by

$$\begin{aligned}
\sum_{\eta \ne 0} [f(\eta) - f(0)]^2 W(\eta) &\le 2 \sum_{\eta \ne 0} \left[f(\eta) - |\bar{\eta}|^{-1} \sum_x \eta_x f(\sigma_x \eta) \right]^2 W(\eta) \\
&\quad + 2 \sum_{\eta \ne 0} \left[|\bar{\eta}|^{-1} \sum_x \eta_x f(\sigma_x \eta) - f(0) \right]^2 W(\eta) \\
&\le 2 \sum_{\eta \ne 0} |\bar{\eta}|^{-1} \sum_x \eta_x [f(\eta) - f(\sigma_x \eta)]^2 W(\eta) \\
&\quad + 2 \sum_{\eta \ne 0} |\bar{\eta}|^{-1} \sum_x \eta_x [f(\sigma_x \eta) - f(0)]^2 W(\eta).
\end{aligned} \tag{5.27}$$

For each x fixed, change the variable $\sigma_x \eta = \zeta$. Hence

$$\begin{aligned}
\sum_{\eta \ne 0} [f(\eta) - f(0)]^2 W(\eta) &\le 2 \sum_\zeta (1 + |\bar{\zeta}|)^{-1} \sum_x (\sigma_x \zeta)_x [f(\sigma_x \zeta) - f(\zeta)]^2 W(\sigma_x \zeta) \\
&\quad + 2 \sum_\zeta (1 + |\bar{\zeta}|)^{-1} \sum_x (\sigma_x \zeta)_x [f(\zeta) - f(0)]^2 W(\sigma_x \zeta) \\
&\le 2\alpha \sum_\zeta \sum_x [f(\sigma_x \zeta) - f(\zeta)]^2 W(\zeta) + 2\alpha |\Gamma| \sum_{\zeta \ne 0} [f(\zeta) - f(0)]^2 W(\zeta).
\end{aligned}$$

Here we have used (5.24) in the last step. By assumption (5.24), $\alpha |\Gamma| < 1/4$, so we have

$$\sum_{\eta \ne 0} [f(\eta) - f(0)]^2 W(\eta) \le 4\alpha \mathscr{D}_\Gamma(f).$$

Together with (5.26) we have proved Lemma 5.4. $\square$

Lemma 5.5. *Let* $V_\varrho(\zeta)$ *denote the marginal distribution of* ν_N *on* Γ *as in Lemma 5.2. Suppose that*

$$|\Gamma| < L^{1/20}, \quad \varrho < L^{-1/8}.$$

Then V_ϱ *satisfies the assumption (5.24) with* $\alpha = $ const. ϱ *and thus it has a spectral gap at least* const.ϱ^{-1}.

Proof. Lemma 5.5 is a simple corollary of Lemma 3.3. For example (5.24) asserts that the probability to have $\eta_x = 1$ is smaller than the probability of having $\eta_x = 0$ by a factor const. ϱ. Since the density $\varrho < L^{-1/8} \ll |\Gamma|^{-1}$, one can follow the same argument as in proving (3.10). We omit the details. □

Corollary 5.6 *Suppose Assumptions A.1 and A.2 hold. Then with the above notation* V *has a spectral gap at least* const. ϱ^{-1} *provided that* $|\Gamma| < L^{1/20}$.

VI. Proof of Theorem 3

The proof of Theorem 3 is very similar to that of Theorem 1. We shall follow the same notation and give details only to those requiring different arguments.

Proof of Theorem 3. Step 1. Instead of the identity (2.2) we use

$$S(f) = \sum_{j=0}^{\infty} E\, E[f_j \log(f_j/f_{j+1}) \mid \mathscr{F}_{j+1}]. \tag{6.1}$$

Steps 2 and 3. Clearly,

$$E[f_j \log f_j/f_{j+1} \mid \mathscr{F}_{j+1}]$$
$$\leq \text{const.} \{-\log[\varrho(1-\varrho)]\}\, \varrho(1-\varrho)E\big[(\sigma_j\sqrt{f_j})^2 \mid \mathscr{F}_{j+1}\big]. \tag{6.2}$$

By definition of $\widetilde{f}_j$ and f_j in (2.5), one has

$$\text{const.}\,\widetilde{f}_j \leq f_j \leq \text{const.}\,\widetilde{f}_j.$$

Together with (2.5),

$$\left|\sqrt{f_j} - \sqrt{\widetilde{f}_j}\right| \leq \text{const.}\,\big|\langle f; h^{(j)}\rangle_{\nu(j)}\big|\widetilde{f}_j^{-1/2}.$$

Hence

$$\varrho(1-\varrho)E[\sigma_j\sqrt{f_j})^2]$$
$$\leq 2\varrho(1-\varrho)E\big[(\sigma_j\sqrt{\widetilde{f}_j})^2\big]$$
$$\quad + 2\varrho(1-\varrho)E[\langle f(\eta_j = 1); h^{(j)}\rangle_{\nu(j)}^2\widetilde{f}_j^{-1}(\eta_j = 1)$$
$$\quad + \langle f(\eta_j = 0); h^{(j)}\rangle_{\nu(j)}^2\widetilde{f}_j^{-1}(\eta_j = 0)]$$
$$\leq \text{const.}\,\varrho(1-\varrho)E\big[(\sigma_j\sqrt{f})^2\big] + \text{const.}\,E\big[\langle f; h^{(j)}\rangle_{\nu(j)}^2\widetilde{f}_j^{-1}\big]. \tag{6.3}$$

Here $h^{(j)}$ and $\nu^{(j)}$ are defined in (2.4), and we have used the bound

$$E\big[(\sigma_j\sqrt{\widetilde{f}})^2\big] \leq \text{const.}\,E\big[(\sigma_j\sqrt{f})^2\big].$$

To prove the last inequality, let z_0^2 and z_1^2 denote $f(\eta_j = 0)$ and $f(\eta_j = 1)$ respectively, then it suffices to prove that

$$E\left[\sqrt{\int z_1^2 d\mu} - \sqrt{\int z_0^2 d\mu}\right]^2 \leq \text{const. } E\left[\int (z_1 - z_0)^2 d\mu\right].$$

But this nothing but the triangular inequality applies to the L^2-norm w.r.t. the measure $d\mu$.

Step 4. Equation (2.13) needs some modification. Consider the variational problem

$$u(\beta) = \sup_{\int f \, d\sigma = 1} \left\{ \int f\left(h_\alpha^{(j)} - h_{\alpha+1}^{(j)}\right) d\sigma - \beta S_\sigma \right\}.$$

Here $\int f \, d\sigma = E[f \mid \mathscr{G}_{\alpha+1}]$ and S_σ denote the entropy with respect to the measure $d\sigma$. By Assumption (A1)

$$\|h_\alpha^{(j)} - h_{a+1}^{(j)}\|_\infty \leq \exp[-C\ell^\alpha]\{\varrho(1 - \varrho)\}^{1/2}. \tag{6.4}$$

By the entropy inequality

$$\int fX \, d\sigma \leq \beta \log \int \exp[\beta^{-1}X] + \beta S_\sigma(f),$$

we have that $u(\beta)$ is bounded by

$$u(\beta) \leq \beta \log \int \exp\left[\beta^{-1}\left(h_\alpha^{(j)} - h_{\alpha+1}^{(j)}\right)\right] d\sigma. \tag{6.5}$$

Suppose that $\beta^{-1} \leq 1$. Then (6.5) is bounded by expanding the exponential to the second order. By definition the first order $\int (h_\alpha^{(j)} - h_{\alpha+1}^{(j)}) \, d\sigma = 0$. Hence $u(\beta)$ is bounded by

$$u(\beta) \leq C\beta^{-1} \exp[-C\ell^\alpha]\{\varrho(1 - \varrho)\}^{1/2}. \tag{6.6}$$

If $\beta^{-1} \geq 1$ then replacing $h_\alpha^{(j)} - h_\alpha^{(j+1)}$ by its maximum $\exp[-C\ell^\alpha]\{\varrho(1 - \varrho)\}^{1/2}$ we have (6.6) holds trivially. Hence by optimizing β

$$\left[\int f\left(h_\alpha^{(j)} - h_{\alpha+1}^{(j)}\right) d\sigma\right]^2 \leq [C\beta^{-1} \exp[-C\ell^\alpha]\{\varrho(1 - \varrho)\}^{1/2} + \beta S_\sigma(f)]^2$$

$$\leq 4CS_\sigma(f) \exp[-C\ell^\alpha]\{\varrho(1 - \varrho)\}^{1/4} \tag{6.7}$$

for all f normalized to $\int f \, d\sigma = 1$. We are now ready to bound $\langle f; h^{(j)} \rangle_{\nu(j)}^2$,

$$\langle f; h \rangle_{\nu(j)}^2$$

$$\leq \sum_{\alpha=0}^\infty \langle f; h_\alpha^{(j)} - h_{\alpha+1}^{(j)} \rangle_{\nu(j)}^2 \alpha^2$$

$$= \text{const.} \sum_{\alpha=0}^\infty \left\{ E^{\nu(j)} E^{\nu(j)}\left[f; h_\alpha^{(j)} - h_{\alpha+1}^{(j)} \mid \mathscr{G}_{\alpha+1}\right] \right\}^2 (\alpha + 1)^2$$

$$\leq \text{const.} \sum_\alpha \left\{ E^{\nu(j)} S_\alpha\left[f(\widetilde{f}_j^{(\alpha)})^{-1}\right]^{1/2} \widetilde{f}_j^{(\alpha)} \right\}^2 \exp\left[-\frac{1}{4}C\ell^\alpha\right]\{\varrho(1 - \varrho)\}^{1/4}, \tag{6.8}$$

where $\widetilde{f}_j^{(\alpha)} = E^{\nu(j)}[f \mid \mathscr{S}_{\alpha+1}]$ and S_α denote the entropy with respect to $E^{\nu(j)}[\mid \mathscr{S}_{\alpha+1}]$. Note that we divide f by $\widetilde{f}_j^{(\alpha)}$ to have a normalized probability density.

Step 5. Again we divide the summation into three regions. We shall only consider the region $\alpha_0 < \alpha < \alpha_1$. By induction

$$S_\alpha\left[f\left(\widetilde{f}_j^{(\alpha)}\right)^{-1}\right] \leq u(C\log L)\mathscr{D}_\alpha\left(\sqrt{f}\right)/\widetilde{f}_j^{(\alpha)}.$$

Here the Dirichlet form $\mathscr{D}_\alpha$ is with respect to $E^{\nu(j)}[\mid \mathscr{S}_{\alpha+1}]$. So we have

$$E\left[\left\{E^{\nu(j)}S_\alpha\left[f\left(\widetilde{f}_j^{(\alpha)}\right)^{-1}\right]^{1/2}\widetilde{f}_j^{(\alpha)}\right\}^2\widetilde{f}_j^{-1}\right]$$

$$\leq \text{const.}\, u(C\log L)E\left\{\left[E^{\nu(j)}\mathscr{D}_\alpha\left(\sqrt{f}\right)^{1/2}\left(\widetilde{f}_j^{(\alpha)}\right)^{1/2}\right]^2\widetilde{f}_j^{-1}\right\}$$

$$\leq \text{const.}\, u(C\log L)E\left[\left\{E^{\nu(j)}\mathscr{D}_\alpha\left(\sqrt{f}\right)\right\}\left\{E^{\nu(j)}\widetilde{f}_j^{(\alpha)}\right\}\widetilde{f}_j^{-1}\right]$$

$$= \text{const.}\, u(C\log L)E\left\{\sum_{x\in\widetilde{B}_\alpha}\left(\sigma_x\sqrt{f}\right)^2\right\}.$$

Conclusion. We can now follow the remaining arguments in Steps 5, 6 and 7 in the proof of Theorem 1 to conclude Theorem 3. □

References

[AZ]   Aizenman, M., Holley, R.: Rapid convergence to equilibrium of the stochastic Ising model in the Dobrushin-Shlosman regime. IMS Vol. in Math. & Appl., Berlin, Heidelberg, New-York: Springer 1987, pp. 1–11

[CS]   Carlen, E. Stroock, D.: An application of the Bakry-Emery criterion to infinite dimensional diffusions. In: Séminaire de Probabilités XX. Springer Lecture Notes in Mathematics **1204**, 1986, pp. 341–347

[DGS] Davies, E.B., Gross, L., Simon, B.: Hypercontractivity: A bibliographical review. In: Ideas and Methods of Mathematics and Physics. In Memoriam of Raphael Hoegh-Krohn. S. Albeverio, J.E. Fenstand, H. Holden, T. Lindstrom (eds.) Cambridge: Cambridge University Press, 1992

[DP]   De Masi A., Presutti, E.: Mathematical Methods for Hydrodynamic Limit. Lecture Notes in Math. **1501**, Berlin, Heidelberg, New York: Springer 1991

[F]   Funaki: private communication

[G]   Gross, L.: Logarithmic Sobolev Inequalities. Am. J. Math. **97**, 1061–1083 (1976)

[KLO] Kipnis, C., Landim, C., Olla, S.: Hydrodynamical limit for a nongradient system: The generalized symmetric exclusion process. Preprint

[MO1] Martinelli, F., Olivieri, E.: Approach to equilibrium of Glauber dynamics in the one phase region, I: The attractive case. Preprint

[MO2] Martinelli, F., Olivieri, E.: Finite volume mixing conditions for lattice spin systems and exponential approach to equilibrium of Glauber dynamics. Preprint

[O]   Olivieri, E.: On a cluster expansion for lattice spin systems with a finite size condition for the convergence. J. Stat. Phys. **50**, 1179–1200 (1988)

[OP]   Olivieri, E., Picco, P.: Cluster expansion for a d-dimensional lattice system and finite volume factorization properties. J. Stat. Phys. **59**, 221–256 (1990)
[Q]   Quastel, J.: Diffusion of color in the simple exclusion process, Thesis, New York Univ. 1990
[R]   Ruelle, D.: Statistical Mechanics: Rigorous Results. New York: Addison-Wesley 1989
[S]   Spohn, H.: Large Scale Dynamics of Interacting Particles. Berlin, Heidelberg, New York: Springer 1991
[ST]   Stroock, D.W.: Logarithmic Sobolev Inequalities for Gibbs States. MIT preprint
[SZ1]   Stroock, D., Zegarlinski, B.: The logarithmic Sobolev inequality for lattice gases with continuous spins. J. Funct. Anal. **104**, 299–326 (1992)
[SZ2]   Stroock, D., Zegarlinski, B.: The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition. Commun. Math. Phys. **144**, 303–323 (1992)
[SZ3]   Stroock, D., Zegarlinski, B.: The Logarithmic Sobolev Inequality for Discrete Spin Systems on a Lattice. Commun. Math. Phys. **149**, 175–193 (1992)
[V]   Varadhan, S.R.S.: Nonlinear diffusion limit for a system with nearest neighbor interactions II. Proc. Taniguchi Symp., 1990 Kyoto
[Z1]   Zegarlinski, B.: private communication
[Z2]   Zegarlinski, B.: Dobrushin Uniqueness Theorem and Logarithmic Sobolev Inequalities. J. Funct. Anal. **105**, 77–111 (1992)
[Z3]   Zegarlinski, B.: Gibbsian Description and Description by Stochastic Dynamics in the Statistical Mechanics of Lattice Spin Systems with Finite Range Interactions. In: Proc. of IIIrd International Conference: Stochastic Processes, Physics and Geometry, Locarno, Switzerland, June 24–29, 1991

Note added in proof. D. Stroock has remarked that the pathological cubes mentioned in the introduction has never appeared in the proof of [SZ2, ST], though the statement of the equivalence theorem is for arbitrary domain.

Communicated by J. L. Lebowitz