**Logarithmic Sobolev Inequality for Lattice Gases**

**with Mixing Conditions**

Horng-Tzer Yau*

Courant Institute, New York University

New York, NY 10012

e-mail: yau@math.nyu.edu

5/1/95

**Abstract**

Let $\mu^{gc}_{\Lambda_L,\lambda}$ denote the grand canonical Gibbs measure of a lattice gas in a cube of size $L$ with the chemical potential $\lambda$ and a fixed boundary condition. Let $\mu^{c}_{\Lambda_L,n}$ be the corresponding canonical measure defined by conditioning $\mu^{gc}_{\Lambda_L,\lambda}$ on $\sum_{x\in\Lambda}\eta_x = n$. Consider the lattice gas dynamics for which each particle performs random walk with rates depending on near-by particles. The rates are chosen such that, for every $n$ and $L$ fixed, $\mu^{c}_{\Lambda_L,n}$ is a reversible measure. Suppose that the Dobrushin-Shlosman mixing conditions holds for $\mu_{L,\lambda}$ for *all* chemical potentials $\lambda \in \boldsymbol{R}$. We prove that $\int f \log f d\mu^{c}_{\Lambda_L,n} \leq \text{const.} L^2 D(\sqrt{f})$ for any probability density $f$ with respect to $\mu^{c}_{\Lambda_L,n}$; here the constant is independent of $n$ or $L$ and $D$ denotes the Dirichlet form of the dynamics. The dependence on $L$ is optimal.

*Keywords*: Dobrushin-Shlosman mixing conditions, Interacting random walks, Lattice gas dynamics, Logarithmic Sobolev inequality

* Research partially supported by U. S. National Science Foundation grant 9403462, Sloan Foundation Fellowship and David and Lucile Packard Foundation Fellowship.

## I Introduction

Suppose that $\mathcal{L}$ is the generator of a dynamics and that $\mu$ is an invariant measure. The Dirichlet form of a function $g$ is defined by

$$D(g) = -\int g\ \mathcal{L}g\ d\mu\,.$$

As only the symmetric part of the generator enters in this definition, we may as well assume that the dynamics is reversible, i.e. $\mathcal{L}$ is symmetric with respect to $\mu$. A logarithmic Sobolev inequality for this system states that the entropy of a probability density $f$ with respect to $\mu$ can be bounded by a constant multiple of the Dirichlet form, namely,

$$\int f \log f d\mu \leq \kappa D(\sqrt{f})$$

It is well-known that the logarithmic Sobolev inequality is equivalent to the hypercontractivity of the semigroup and thus it provides certain information on the relaxation to equilibrium for the dynamics [G, DGS, DS, D]. In particular, it implies that the spectral gap of the generator, and hence the relaxation rate in the $L^2$ norm, is bounded by a constant multiple of $\kappa^{-1}$. See [DS, DGS] for references and historical remarks. Indeed, from the logarithmic Sobolev inequality one can obtain that the relaxation rate in certain semi-norm much stronger than the $L^2$ norm is bounded by a constant multiple of $\kappa^{-1}$ provided some mild conditions on the systems are given [HS,S, SZ]. See [S] for a recent review.

In this article, we will estimate the constant $\kappa$ for lattice gases. The lattice gases can be described as follows. Let $\Lambda$ be a cube of width $L$ in $\mathbb{Z}^d$. At each lattice site of $\Lambda$, we associate an occupation number of particle $\eta_x \in \{0, 1\}$. The equilibrium states of lattice gases are described by the Gibbs measures on $\Lambda$, characterized by a Hamiltonian and a boundary condition. There are two "ensembles" of interest: the grand canonical ensemble with the chemical potential specified and the canonical ensemble with the total number of particles specified. The first measure is denoted by $\mu^{gc}_{L,\lambda}$ where $\lambda$ is the chemical potential; the second measure is denoted by $\mu^{c}_{L,n}$, where $n$ is the total number of particles.

The Hamiltonian and the boundary condition are fixed and will not be specified in the notations.

The dynamics of lattice gases is determined as follows. Each particle performs a random walk with jump rates determined by nearby particles according to some fixed local rules such that the Gibbs measures are reversible measures. To maintain the requirement of at most one particle per site, jumps to occupied sites are suppressed. Because that no creation or annihilation of particles is allowed, the total number of particles is conserved by the dynamics. Therefore, the natural ensemble for this dynamics is the canonical ensemble. The models we have just described are often referred to as symmetric simple exclusion processes with speed change, Kawasaki dynamics or simply lattice gas dynamics. They are systems of interacting random walks, and have a natural interpretation as discretizations of interacting Brownian motions. The simplest example of lattice gases is the well known symmetric simple exclusion process. The dynamics is given by the usual symmetric random walk and the invariant measures are simply product of Bernoulli measures. Except this special case, the jump rates of particles depend on the environments of the particles.

The main result of this paper states that

$$\int f \log f d\mu^c_{L,n} \le \kappa_{L,n} D(\sqrt{f}), \quad D(g) = -\int g \; \mathcal{L} g \; d\mu^c_{L,n} \,. \tag{1.1}$$

with

$$\kappa_{L,n} \le \; \text{const.} L^2, \tag{1.2}$$

for some constant independent of $n$ or $L$. It is easy to check that $\kappa_{L,n} \ge CL^2$ by using test functions. Hence (1.2) identifies the dependence of $\kappa_{L,n}$ on $L$. We do not know as yet the dependence of $\kappa_{L,n}$ on the number of particles $n$.

It is well-known [DGS, G] that the LSI implies a bound on the spectral gap. Hence the result of [LY] on the spectral gap of lattice gases is a direct consequence of this paper. In the special case when there is only one particle, the dynamics is just the usual random walk on $\Lambda$; the spectral gap and the LSI can be computed explicitly. They become the familiar spectral gap and logarithmic Sobolev inequality of the discrete Laplacian. The other special case is the symmetric simple exclusion process. This model is no longer exactly computable, but can be solved almost exactly using duality. It is proved that the gap behaves like const.$L^{-2}$ by, among others, [Q]. The LSI is harder and is proved in [Y] with the correct order $\kappa_{L,n} \leq CL^2$. It is also announced that $\kappa_{L,n} \leq L^2 \log L$ in [DSa,F]. The result in [Y] is valid provided the number of particles per site is finite.

Recently, large numbers of articles on the logarithmic Sobolev inequality have appeared. Apart from those mentioned above, systematic studies of Glauber dynamics of lattice gases were done by, e.g., [HS, Z, SZ, MO, LY]. The Glauber dynamics is a dynamics without conservation law. It prescribes local rules to create or annihilate particles and thus the total number of particles is not conserved. For this dynamics it was proved that there is a uniform LSI independent of the volume of the cube [HS,Z, SZ, MO, LY]; this implies [HS,SZ] the exponential convergence to equilibrium in a certain semi-norm for the infinite volume dynamics. For this dynamics, a local disturbance is expected to stay local and relaxes to equilibrium exponentially fast. For the lattice gases, because $\kappa_{L,n} \sim L^2$, no uniform exponential relaxation is allowed. Indeed, one expects a power law decay in infinite volume. Even the equilibrium (truncated) correlation function, $\langle \eta_x; \eta_y \rangle = \langle \eta_x \eta_y \rangle - \langle \eta_x \rangle \langle \eta_y \rangle$, $x \neq y$, displays different properties. Restricted to the high temperature region, we have, $\langle \eta_x; \eta_y \rangle_{\mu^{gc}_{\Lambda_L,\lambda}} \sim \exp[-\text{const.}|x-y|]$. For the canonical ensemble, even in the infinite temperature case (i.e., product of Bernoulli measures with total number of particles fixed), we have $\langle \eta_x; \eta_y \rangle_{\mu^c_{L,n}} \sim -L^{-d}$. The last estimate on the correlation function is due in part to the conservation of the total number of particles. If the number of particles at a site $x$, say, increases by one, a particle needs to be removed somewhere on the lattice $\Lambda_L$ because of the conservation law. Assuming the probability to remove such a particle is uniform on $\Lambda_L$, we obtain the order of magnitude $L^{-d}$. This negative correlation, albeit small, is in a sense the underlying reason that (1.2) holds.

In field theory terminology, the conservative dynamics is the *massless* case and the nonconservative case is the *massive* case. To deal with massless dynamics, certain *multiscale analysis* and/or the so called *renormalization group* approach are usually needed. Our approach is based on a combination of the martingale method and some ideas from the renormalization group method and the multiscale analysis. Related ideas were used in [LY, Y] but in a more primitive form. The multiscale analysis will be carried out in a manner very different from [LY] or [Y]. The present approach provides a much stronger result. Though our proof is not as simple as we wish due to the use of the multiscale analysis, we believe it is still much simpler than setting up the full renormalization group, should such a approach be mathematically feasible.

Though our approach is quite general, strong mixing conditions on the underlying measures are needed. In [LY], the mixing conditions are summarized as assumptions A.1-A.3. We emphasize that these mixing conditions are w.r.t. *canonical* Gibbs states rather than w.r.t. grand canonical Gibbs states. One expects that the mixing conditions w.r.t. canonical Gibbs states should follow from certain mixing properties of the corresponding grand canonical Gibbs states, however, no results have been proved. In this paper, we simplify these assumptions to a single assumption w.r.t. *grand* canonical Gibbs states, namely, assumption A.1 in Sect. 2. The assumption A.1 holds in particular for ferromagnetic Ising models up to the critical point [MOS, N] in dimension $d = 2$. Certainly, because the total number of particles is conserved by the lattice gases dynamics, some mixing properties w.r.t. canonical Gibbs states are needed. These properties will be proved as consequences of assumption (A.1). Our methods can be used to give a rigorous derivation of the mixing conditions in [LY] from the assumption A.1. We shall not carry this out here because [LY] is a direct corollary of present paper.

The organization of this paper is as follows. In Sect.2 we state the main results. Sect.3 contains an outline of the martingale approach. Sect.4-Sect.8 contain proofs of results assumed in Sect.3. The large deviation estimates needed in this paper will be presented in Sect. 5; the multiscale analysis will be presented in sect. 6. Finally, we prove a version of local limit theorem for the Gibbs states with mixing conditions in Sect.9.

## II Statement of Main Results

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

$$\partial\Lambda = \{y \in \mathbb{Z}^d \setminus \Lambda \mid \mathrm{dist}(y, \Lambda) = 1\} \ , \tag{2.1}$$

where the distance function is defined by

$$\begin{aligned} \mathrm{dist}(y, \Lambda) &= \inf_{x \in \Lambda} |x - y| \ , \\ |x - y| &= \max_{\alpha=1,\ldots,d} |x^\alpha - y^\alpha| \ . \end{aligned} \tag{2.2}$$

Let $\omega$ be a configuration on $\partial\Lambda$ with $\omega_x$ belonging 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

$$Z_p = \{0, 1, 2, \cdots, p-1\} \ , \qquad 2 \le p \in N \ . \tag{2.3}$$

The Hamiltonians are finite ranged and translationally invariant. For simplicity of notation, we restrict ourselves to nearest neighbor interactions. All our results hold for Hamiltonian with finite range interactions. Thus the Hamiltonian is characterized by an interaction $J(\eta_x, \eta_y)$ such that

$$H_{\Lambda,\omega}(\eta) = \sum_{x,y \in \Lambda, |x-y|=1} J(\eta_x, \eta_y) + \sum_{y \in \partial\Lambda,\ x \in \Lambda,\ |x-y|=1} J(\eta_x, \omega_y) \ . \tag{2.4}$$

The grand canonical Gibbs state with the chemical potential $\lambda$ and the boundary condition $\omega$ is characterized by the density

$$d\mu^{gc}_{\Lambda,\omega,\lambda}(\eta) = \exp[-H(\eta) + \lambda \sum_{x \in \Lambda} \eta_x] / Z^{gc}_{\Lambda,\omega,\lambda} \ . \tag{2.5}$$

Here the partition function $Z^{gc}_{\Lambda,\omega,\lambda}$ is the normalization factor to make $d\mu^{gc}_{\Lambda,\omega,\lambda}$ into a probability density. We shall denote by $E^{\mu^{gc}_{\Lambda,\omega,\lambda}}$ or $\langle\ \rangle_{\mu^{gc}_{\Lambda,\omega,\lambda}}$ the expectation with respect to $d\mu^{gc}_{\Lambda,\omega,\lambda}$. When $\lambda = 0$, we shall drop the subscript $\lambda$.

We need the concept of canonical Gibbs states. Let $n$ be a fixed positive integer. A canonical Gibbs state with total number of particles $n$ and boundary condition $\omega$ is characterized by the density

$$d\mu^c_{\Lambda,\omega,n} = d\mu^{gc}_{\Lambda,\omega,\lambda}\Big|_{|\Lambda|\bar\eta=n} \tag{2.6}$$

Here $\bar\eta = |\Lambda|^{-1}\sum_{x\in\Lambda}\eta_x$ is the density. Note that the right side of (2.6) is independent of $\lambda$. Define also the canonical partition function

$$Z^c_{\Lambda,\omega,n} = \sum_{\bar\eta=n}\exp[-H_{\Lambda,\omega}(\eta)] \ . \tag{2.7}$$

We shall drop the subscript $\omega$ if the boundary condition plays no active role. Also, in this section $\Lambda$ denotes a cube of width $L$, i.e., $\Lambda=\Lambda_L$. We also denote $d\mu^c_{\Lambda_L,\omega,n}$ by $d\mu^c_{L,\omega,n}$.

For any function $g$ on the configuration space, define two operators $\sigma_x g(\eta) = g(\sigma_x\eta)$ and $T_{xy}g(\eta) = g(T_{xy}\eta)$. Here $\sigma_x\eta = \eta^x$ and $T_{xy}\eta = \eta^{xy}(x\neq y)$ are defined by

$$(\sigma_x\eta)_y = (\eta^x)_y = \delta_{xy}(1-\eta_y) + (1-\delta_{xy})\eta_y \ . \tag{2.8}$$

$$(T_{x,y}\eta)_z := (\eta^{x,y})_z = \begin{cases} 0, & \text{if } z=x, \eta_x=1 \text{ and } \eta_y=0 \ , \\ 1, & \text{if } z=y, \eta_x=1 \text{ and } \eta_y=0 \ , \\ \eta_z & \text{otherwise} \end{cases} \tag{2.9}$$

From the definition, $T_{xy}\eta$ denote the configuration obtained by moving a particle from $x$ to $y$. Define also the symmetrization of $T_{xy}$ by

$$\tilde T_{xy} = T_{xy} + T_{yx} \ . \tag{2.10}$$

Then the Dirichlet form of the bond $b=(x,y)$ is defined by

$$D_{xy}(h) = \int \left(\tilde T_{xy}h - h\right)^2 d\mu^c_{\Lambda,\omega,n} \tag{2.11}$$

To state our main result, we need the following assumptions. Define, for two boundary conditions $\omega_1$ and $\omega_2$, the set

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

Though our goal is to prove a LSI for standard cubes, in the proof we shall encounter, for example, in the case of dimension $d = 2$, rectangles of sizes $L_1 \times L_2$ with $L_i \leq 2L$ and $1 \leq L_1/L_2 \leq 2$. Similarly for dimension $d \geq 3$. We shall not distinguish such rectangles from standard cubes and we will refer them as "cubes" . Furthermore, all proofs will be carried out only for standard cubes but will be used freely for rectangles as well. It should be noted that in some cases our results can fail if the rectangles degenerate, for example, if they become very "thin", i.e., the length of one side becomes very small [MO].

**Assumption A1** *Let $g$ be a function depending only on the configuration of a subset $U$ in a cube $\Lambda$ of size $L$. Then*

$$\left| E^{\mu^{gc}_{\Lambda,\omega_1,\lambda}}[g] - E^{\mu^{gc}_{\Lambda,\omega_2,\lambda}}[g] \right| \leq C(g)|A_{\omega_1,\omega_2}| \exp[-const.\ dist(A_{\omega_1,\omega_2}, U)]\ . \tag{2.12}$$

*Here the constant $C(g)$ is independent of $\Lambda$ and $\omega_i, i = 1, 2$.*

**Theorem 2.1.** *Suppose that the Gibbs measures satisfy the mixing condition (A1) for all $\lambda$. Let $\Lambda$ be a cube of width $L$. Then there is a constant $C$ independent of $L$, $n$ such that for any probability density function $f$ (i.e. $\int f d\mu^c_{\Lambda,\omega,n} = 1$)*

$$E^{\mu^c_{\Lambda,\omega,n}}[f \log f] \leq CL^2 D_\Lambda(\sqrt{f}). \tag{2.13}$$

*where*

$$D_\Lambda(g) := \sum_{b\in\Lambda} D_b(g) := \sum_{b\in\Lambda} \int \left[ \left( \tilde{T}_{x,y} g - g \right)^2 \right] d\mu^c_{\Lambda,\omega,n}, \quad b = (x, y),\ |x - y| = 1$$

From the Dirichlet form, we can recover the generator of the dynamics. The generator plays no role in this paper and we shall concentrate on the Dirichlet form $D_\Lambda$. The mixing assumption (A1) can be weakened somehow. Furthermore, (A1) is a consequence of the Dobrushin-Shlosman mixing conditions. It should be emphasized that the mixing condition (A.1) is with respect to grand canonical Gibbs states for *all* $\lambda$ rather than with respect to the canonical Gibbs state $\mu^c_{\Lambda,\omega,n}$, which is the underlying measure in (2.13). Assumption (A1) can be checked, in general, by the high temperature expansion. If lattice gases are described by the ferromagnetic Ising model, the mixing condition holds up to the critical temperature in dimension $d = 2$ [MS, N].

Since we need mixing conditions for all chemical potentials $\lambda$, our result is also uniform w.r.t. the particle numbers. Hence it excludes an interesting case when the pair $(\rho, T)$ for the density and temperature is in the one phase region but $(\rho', T)$ is in the phase transition region for some choice of $\rho'$. We do not know whether the L.S.I. has very different prefactor in this case. Unlike the case of grand canonical ensemble, it is relatively easy for the canonical ensemble to have some region with density very different from the global density $\rho$ and this may change the prefactor for the L.S.I. Unfortunately, we do not have rigorous result along this direction.

Because Theorem 2.1 concerns canonical ensembles, some mixing properties with respect to canonical Gibbs states will be needed in order to prove Theorem 2.1. For the convenience of later references we list them as (A2-A4). *We shall prove Theorem 2.1 assuming (A1-A4)* . This will be done in Sect.III-Sect.VII. A derivation of (A2-A4) from (A1), stated as Theorem 2.2, will be proved in Sect.VIII and Sect.IX.

In the following (A2-A4), $\Lambda$ is a cube of size $L$.

*(A2) For any local function $g$ let*

$$\hat{g}(y) = E^{\mu^{gc}_{\Lambda,\omega,\lambda}}[g]$$

*where $\lambda$ is chosen so that the density is $y$. Then $\partial\hat{g}/\partial y \le const.$ and $\partial^2\hat{g}/\partial y^2 \le const.$ with a constant independent of $\lambda$, $\omega$ or $L$.*

*(A3) Suppose $x$ and $z$ are two nearest neighbor points in $\Lambda$ with $|x - \partial\Lambda| \ge \sigma L$ for some small constant $\sigma$. Then for some small constant $\epsilon$,*

$$\| E^{\mu^c_{\Lambda,\omega,n}}[\eta_x - \eta_z] \|_\infty \le const. L^{-(d/2)-1-\epsilon}$$

*uniformly with respect to the boundary condition and $n$.*

*(A4) Suppose that $\mu^{gc}_{\Lambda,\omega,\lambda}$ and $\mu^{c}_{\Lambda,\omega,n}$ are grand canonical and canonical Gibbs states with the same boundary conditions on $\Lambda$ and $E^{\mu^{gc}_{\Lambda,\omega,\lambda}}[\bar\eta] = n/|\Lambda|$ . Then there is an $\epsilon > 0$ such that for any local function $g$*

$$|E^{\mu^{gc}_{\Lambda,\omega,\lambda}}[g] - E^{\mu^{c}_{\Lambda,\omega,n}}[g]| \leq const. L^{-d}$$

*uniformly with respect to the boundary condition, $\lambda$ and $n$. Here the constant may depend on $g$.*

**Theorem 2.2.** *Suppose the Gibbs measures satisfy the mixing condition (A1) for all $\lambda$. Then (A2-A4) hold.*

Finally we remark on a convention of notations. We shall drop all superscripts $gc$ and $c$. To distinguish them, for grand canonical ensmebles we have a chemical potential subscript $\lambda$; for canonical ensmebles we have a total number of particles subscript $n$.

## III Outline of Martingale Approach

In this section we outline the martingale approach of [LY]. This section is almost identical to the same section in [Y], where the LSI for the independent random variables (with the total number of particles fixed) is proved. All difficulties related to interactions will appear in the proof of Theorem 3.1. For simplicity of notation, we shall assume $d = 2$ unless otherwise noted. Assume that Theorem 2.1 holds for $\Lambda$ of size $L \times L$. Our goal is to prove that Theorem 2.1 holds for $\Lambda$ of size $L \times 2L$ and then for $\Lambda$ of size $2L \times 2L$. As the proofs from $L \times L$ to $L^2$ and from $L \times 2L$ to $2L \times 2L$ are identical, we will only prove Theorem 2.1 for $\Lambda$ of size $L \times 2L$ assuming it holds for $\Lambda$ of size $L \times L$. Let us order the sites in the lower half of $\Lambda_{L\times 2L}$ lexicographically by $i = 1, 2, \dots, L \times 2L$ and denote the upper half of $\Lambda$ by $\Lambda_0$. Define $u(2L)$ to be the smallest constant such that for any cube $\Lambda$ of size $L_1 \times L_2$ with $L_i \leq 2L$ and $1 \leq L_1/L_2 \leq 2$ and any probability density $f$

$$E^{\mu_{\Lambda,n}}[f \log f] \leq u(2L)(2L)^2 D_\Lambda(\sqrt{f}). \tag{3.1}$$

**Proof of Theorem 2.1.** *Step 1.* Denote by $\mathcal{F}_j$ the $\sigma$ - algebra generated by $\eta_j, \eta_{j+1}, \cdots$. Define the marginal density $f_j$ by

$$f_j(\eta_j, \eta_{j+1} \dots) = E^{\mu_{L\times 2L,n}}[f|\mathcal{F}_j],\ j \geq 1\,;\ f_0 = f\,.$$

Here $E^{\mu_{L\times 2L,n}}$ is the expectation with respect to the Gibbs state on $\Lambda_{L\times 2L}$. Then one has the identity

$$E^{\mu_{L\times 2L,n}}[f\log f] = \sum_{j=0}^{\infty} E^{\mu_{L\times 2L,n}} E^{\mu_{L\times 2L,n}}[f_j \log(f_j/f_{j+1})|\mathcal{F}_{j+1}]\ . \tag{3.2}$$

Note that the summation terminates at $j = L^2 - 1$. By the inductive hypothesis,

$$E^{\mu_{L\times 2L,n}}\left\{E^{\mu_{L\times 2L,n}}[f\log(f/f_1)|\mathcal{F}_1]\right\} \leq u(L)L^2 \sum_{b\in\Lambda_0} D_b(\sqrt{f}) = u(L)L^2 D_{\Lambda_0}(\sqrt{f}) \tag{3.3}$$

where $b$ denotes a typical bond in $\Lambda_0$.

*Step 2.* We now bound the right side of (3.2) for $j > 0$. Our main result can be stated as the following Theorem 3.1 which will be proved in Sect. IV-VII.

**Theorem 3.1** *Let $\Lambda$ be a cube of size $L_1 \times L_2$ with $1 \leq L_i/L \leq 2$ for $i = 1, 2$. Assume the mixing condition (A1-A4) hold. For any probability density $f$ with respect to $\mu_{\Lambda,n}$ let $\bar{f}_z(\eta_z)$ denote the marginal density of $\eta_z$ for some $z \in \Lambda$. For any path $\gamma_{zy}$ from $z$ to $y$, label the path by $\gamma_i, i = 0, \cdots, |\gamma|$ and define*

$$D^{\gamma_{zy}}(\sqrt{f}) = \sum_{i=1}^{|\gamma|} D_{\gamma_i-1,\gamma_i}(\sqrt{f}), \quad D_{\gamma_i-1,\gamma_i}(g) = \int \left[(\tilde{T}_{\gamma_{i-1},\gamma_i} g - g)^2\right] d\mu_{\Lambda,n}\,. \tag{3.4}$$

*Define also the function*

$$\Psi(v) = \begin{cases} v^{2-d} & \text{if } d \geq 3, \\ v^{2-d}\{\log[L/v]\}^3 & \text{if } d = 2, \\ L\{\log[L/v]\}^3 & \text{if } d = 1. \end{cases} \tag{3.5}$$

*Here $d$ is the dimension of the lattice $\Lambda$ ($d = 2$ in our setting). The path $\gamma_{zy}$ will be chosen in a canonical way, e.g., from $z = (z_1, z_2)$ to $(z_1, y_2)$ and then to $(y_1, y_2) = y$. Then for* *any $\delta > 0$ and integer $\ell$ fixed, there is a constant $C_1$ such that*

$$E^{\mu_{\Lambda,n}}[\bar{f}_z \log \bar{f}_z] \le C_1 L^{1-d} \sum_{y\in\Lambda} D^{\gamma_{zy}}(\sqrt{f}) + C_1 L^{2-d} D_\Lambda(\sqrt{f}) + \delta u(2L) L^{2-d} D_\Lambda(\sqrt{f})$$
$$+ C_1 u(L) \sum_{|z-b|\le L/\ell} \Psi(|z-b|) D_b(\sqrt{f}), \tag{3.6}$$

*where $u(2L)$ is defined in (3.1).*

The form of the last term in (3.6) will be explained in Sect.4 and Sect.5. Applying Theorem 3.1 to bound (3.2) for $j > 0$, we have

$$\sum_{j=1}^{\infty} E^{\mu_{L\times 2L,n}} E^{\mu_{L\times 2L,n}}[f_j \log(f_j/f_{j+1})|\mathcal{F}_{j+1}]$$
$$\le C_1 L^{1-d} \sum_{x\in\Lambda_{L\times 2L}} \sum_{y\in\Lambda_{L\times 2L}} D^{\gamma_{xy}}(\sqrt{f}) + C_1 L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) + \delta u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f})$$
$$+ C_1 u(L) \sum_{z\in\Lambda_{L\times 2L}} \sum_{|b-z|\le L/\ell} \Psi(|z-b|) D_b(\sqrt{f})$$

One can checked that by choosing $\ell$ large (independent of $L$),

$$C_1 u(L) \sum_{z\in\Lambda_{L\times 2L}} \sum_{|b-z|\le L/\ell} \Psi(|z-b|) D_b(\sqrt{f}) \le \delta u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}).$$

Hence

$$\sum_{j=1}^{\infty} E^{\mu_{L\times 2L,n}} E^{\mu_{L\times 2L,n}}[f_j \log(f_j/f_{j+1})|\mathcal{F}_{j+1}]$$
$$\le 2C_1 L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) + 2\delta u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}). \tag{3.7}$$

*Step 3.* Combining (3.2), (3.3) and (3.7), one has

$$E^{\mu_{L\times 2L,n}}[f \log f] \le C_2 L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) + 2\delta u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) + u(L) L^2 D_{\Lambda_0}(\sqrt{f}). \tag{3.8}$$

Switching the role of $\Lambda_0$ and $\Lambda_{L\times 2L}\backslash\Lambda_0$, and taking the average, we have

$$E^{\mu_{L\times 2L,n}}[f \log f] \le [C_2 + u(L)/8 + \delta u(2L)](2L)^2 D_\Lambda(\sqrt{f}). \tag{3.9}$$

We can repeat step 1-3 once more and obtain for $\Lambda$ of size $2L \times 2L$,

$$E^{\mu_{\Lambda_{2L\times 2L},n}}[f \log f] \le [2C_2 + u(L)/2 + 4\delta u(2L)](2L)^2 D_{\Lambda_{2L\times 2L}}(\sqrt{f}). \tag{3.10}$$

Since $\delta$ can be chosen as small as we wish, we have proved, from the definition of $u(2L)$,

$$u(2L) \leq \text{ const.} + 2u(L)/3.$$

From the Gronwal's inequality or induction we have proved Theorem 2.1 assuming Theorem 3.1.

## IV Proof of Theorem 3.1

We now prove Theorem 3.1. Throughout this section, $\Lambda$ denotes a cube specified in Theorem 3.1. We divide the proofs into 4 steps.

*Step 1.* Let $\rho = n/|\Lambda| = E^{\mu_{\Lambda,n}}[\eta_x]$ be the particle density. By definition

$$\bar{f}_z(\eta_z = 1) = E^{\mu_{\Lambda,n}}[f|\eta_z = 1] = E^{\mu_{\Lambda,n}}[f\eta_z]\rho^{-1}, \tag{4.1}$$

where the expectation is taken with respect to $\mu_{\Lambda,n}$. Note that $\bar{f}_z(\eta_z)$ is just a function on a single site. One can check easily that the LSI holds trivially for Bernoulli measures on one site. Thus one has

$$\begin{aligned} s(\bar{f}_z/\bar{\mu}^{(z)}_{\Lambda,n}) &= \rho\bar{f}_z(1)\log\bar{f}_z(1) + (1-\rho)\bar{f}_z(0)\log\bar{f}_z(0) \\ &\leq \text{const.}\min\Big[\rho|\log\rho|, (1-\rho)|\log(1-\rho)|\Big] E^{\mu_{\Lambda,n}}[(\sqrt{\bar{f}_z(1)} - \sqrt{\bar{f}_z(0)})^2]. \end{aligned} \tag{4.2}$$

Here $\bar{\mu}^{(z)}_{\Lambda,n}$ is the marginal of $\mu_{\Lambda,n}$ at $z$. We remark that there is a logarithmic correction appearing in (4.2). *This should not be dismissed as merely a technical factor appearing in the very low or very high density region, since in many applications the LSI is used precisely to control the probability in this region.* Because of the particle-hole duality, we can assume the density $\rho$ is bounded by $1/2$. *We shall assume $\rho \leq 1/2$ for the rest of this paper.* Note that if $\rho \leq 1/2$ and

$$\bar{f}_z(1) \leq 10,\ \bar{f}_z(0) \leq 10 \tag{4.3}$$

then the logarithmic factor can be dropped in (4.2). This can be easily checked or see [Y]for a proof.

*Step2.* The starting point to bound the right side of (4.2) is the following lemma 4.1 from [LY]. Roughly speaking, it states that the commutator of the conditional expectation and differentiation (spin flip at one site) can be bounded by a variance term and an exchange term involving long jumps. The exchange term will be bounded using a lemma from [SY], while the variance term will be bounded using a multiscale analysis. Let us first fix some notation.

Recall that $\mu_{\Lambda,n}$ is a canonical Gibbs state on $\Lambda$ with some fixed boundary condition and total number of particles $n$. Let $z$ be a point in $\Lambda$ and denote configurations in $\Lambda$ by $\eta = (\eta_z, \xi)$. Define

$$H_0^{(z)}(\xi) = H(\eta_z = 0,\ \xi)\ .$$

Denote by $\nu_{\Lambda,n}^{(z)}$ the canonical Gibbs state with Hamiltonian $H_0$ and number of particles $n$. Let $\tilde{H}^{(z)}$ be the difference

$$\tilde{H}^{(z)}(\eta_z, \xi) = H(\eta_z, \xi) - H_0^{(z)}(\xi)\ ,$$

Here $H$ is the Hamiltonian defined in (2.4). For each $x \in \Lambda$ define

$$F_x = (1 - \eta_x) \exp\{-\tilde{H}^{(z)}(\sigma_x \eta) + \tilde{H}^{(z)}(\eta)\}\ .$$

Also define

$$\bar{F} = Av_{x\in\Lambda\setminus\{z\}} F_x = (|\Lambda| - 1)^{-1} \sum_{x\in\Lambda\setminus\{z\}} F_x\ . \tag{4.4}$$

Let $g_z$ be defined by

$$g_z = \exp\left[-\tilde{H}^{(z)}(\eta_z, \xi)\right] \,/\, E^{\nu_{\Lambda,n}^{(z)}}\left[\exp[-\tilde{H}^{(z)}(\eta_z, \xi)] \mid \eta_z\right] = d\mu_{\Lambda,n}\big|_{\eta_z} \,/\, d\nu_{\Lambda,n}^{(z)}\big|_{\eta_z}\ . \tag{4.5}$$

We shall use the symbol $E^{\nu_{\Lambda,n}^{(z)}}[f; g \mid \eta_z = 0]$ to denote the covariance

$$E^{\nu_{\Lambda,n}^{(z)}}[f; g \mid \eta_z = 0] = E^{\nu_{\Lambda,n}^{(z)}}[fg \mid \eta_z = 0] - E^{\nu_{\Lambda,n}^{(z)}}[f \mid \eta_z = 0] E^{\nu_{\Lambda,n}^{(z)}}[g \mid \eta_z = 0]$$

This convention will be used for the rest of this paper. The following lemma 4.1 extends a lemma of [Y] to the interacting case.

**Lemma 4.1.** *Recall the definitions of $T_{xy}$ (2.9). Then with the above notations,*

$$[\sqrt{\bar{f}_z(1)} - \sqrt{\bar{f}_z(0)}]^2 \le const.\ (1-\rho)^2\rho^{-1} Av_x E^{\nu^{(z)}_{\Lambda,n}}[\{T_{zx}\sqrt{f}\}^2 \mid \eta_z = 0] + const.\ \Phi, \quad (4.6)$$

*where*

$$\begin{aligned}\Phi =&[\bar{f}_z(1) + \bar{f}_z(0)]^{-1}\times \\ &\Big\{E^{\nu^{(z)}_{\Lambda,n}}[f;\bar{F} \mid \eta_z = 1]^2 + 4E^{\nu^{(z)}_{\Lambda,n}}[f;g_z \mid \eta_z = 0]^2 + 4E^{\nu^{(z)}_{\Lambda,n}}[f;g_z \mid \eta_z = 1]^2\Big\}.\end{aligned}$$

*Note that the left side of (4.6) is independent of $\eta_z$.*

By definition

$$E^{\nu^{(z)}_{\Lambda,n}}[\{T_{zx}\sqrt{f}\}^2 \mid \eta_z = 0] \le E^{\nu^{(z)}_{\Lambda,n}}[\{T_{zx}\sqrt{f}\}^2(1-\eta_z)]E^{\nu^{(z)}_{\Lambda,n}}[1-\eta_z]^{-1}\ .$$

Since $d\nu^{(z)}_{\Lambda,n}/d\mu_{\Lambda,n}$ is uniformly bounded,

$$E^{\nu^{(z)}_{\Lambda,n}}[\{T_{zx}\sqrt{f}\}^2 \mid \eta_z = 0] \le E^{\mu_{\Lambda,n}}[\{T_{zx}\sqrt{f}\}^2(1-\eta_z)]E^{\mu_{\Lambda,n}}[1-\eta_z]^{-1}\ .$$

We need the following elementary fact.

For any canonical Gibbs measure (with finite range Hamiltonian) $\mu_{\Lambda,n}$ and any $z \in \Lambda$ we have

$$\text{const.} \le E^{\mu_{\Lambda,n}}[\eta_z]/\rho \le \text{const.} \quad (4.7)$$

with the constant independent of $\rho$. To prove this, it suffices to prove that for any two sites $z, y \in \Lambda$,

$$E^{\mu_{\Lambda,n}}[\eta_z] \le \text{const.} E^{\mu_{\Lambda,n}}[\eta_y] \quad (4.8)$$

Assuming this bound, we obtain the upper bound of $E^{\mu_{\Lambda,n}}[\eta_z]/\rho$ by averaging over $y \in \Lambda$. Exchanging the role of $z$ and $y$ and repeating the same procedure, we obtain the lower bound. We now prove (4.8). Note (4.8) holds if the parts of Hamiltonian involving $\eta_z$ or $\eta_x$ are removed. On the other hand, these parts is bounded by a constant. Hence (4.8) holds.

Recall the assumption $\rho \le 1/2$. Hence $E^{\mu_{\Lambda,n}}[1-\eta_z]^{-1} \le$ const.. Combining (4.3) and (4.6) we have

$$s(\bar{f}_z/\bar{\mu}^{(z)}_{\Lambda,n}) \le \text{const.}\ \Big\{|\log\rho| Av_x E^{\mu_{\Lambda,n}}[\{T_{zx}\sqrt{f}\}^2] + \rho|\log\rho|\Phi\Big\}. \quad (4.9)$$

The rest of this step is devoted to a proof of Lemma 4.1.

**Proof of Lemma 4.1.** By definition of $g_z$, $\nu^{(z)}_{\Lambda,n}$ and covariance,

$$\bar{f}_z(\eta_z) = E^{\mu_{\Lambda,n}}[f \mid \eta_z] = E^{\nu^{(z)}_{\Lambda,n}}[fg_z \mid \eta_z] = E^{\nu^{(z)}_{\Lambda,n}}[f; g_z \mid \eta_z] + E^{\nu^{(z)}_{\Lambda,n}}[f \mid \eta_z] \ ,$$

where we have used $E^{\nu^{(z)}_{\Lambda,n}}[g_z \mid \eta_z] = 1$. From the elementary inequality

$$\left[\sqrt{\bar{f}_z(1)} - \sqrt{\bar{f}_z(0)}\right]^2 \leq [\bar{f}_z(1) - \bar{f}_z(0)]^2/[\bar{f}_z(1) + \bar{f}_z(0)],$$

and the Schwarz inequality one has

$$\left[\sqrt{\bar{f}_z(1)} - \sqrt{\bar{f}_z(0)}\right]^2 \leq 2\left\{E^{\nu^{(z)}_{\Lambda,n}}[f; g_z \mid \eta_z = 1]^2 + E^{\nu^{(z)}_{\Lambda,n}}[f; g_z \mid \eta_z = 0]^2\right\} \left[\bar{f}_z(1) + \bar{f}_z(0)\right]^{-1}$$
$$+ 2\left\{E^{\nu^{(z)}_{\Lambda,n}}[f \mid \eta_z = 1]^2 - E^{\nu^{(z)}_{\Lambda,n}}[f \mid \eta_z = 0]^2\right\} \left[\bar{f}_z(1) + \bar{f}_z(0)\right]^{-1} .$$

To conclude Lemma 4.1, it suffices to bound the last term. We need the following Lemma.

**Lemma 4.2** *Suppose $f$ is a function on $U \cup \{0\}$ and $0 \notin U$. Define $\bar{f}(m) = E^{\omega_{U,m}}[f]$ where $\omega_{U,m}$ is a canonical Gibbs measure with boundary condition independent of $\eta_z$. Then*

$$\{\bar{f}(m+1) - \bar{f}(m)\}^2/[\bar{f}(1) + \bar{f}(z)] \leq A_1 + A_2$$
$$A_1 = const.(1 - m/|U|)^{-1} G(m+1) Av_x E^{\omega_{U,m+1}}\left[\{\sqrt{f(T_{x0}\eta)} - \sqrt{f(\eta)}\}^2\right]$$
$$A_2 = const.(1 - m/|U|)^{-2} E^{\omega_{U,m}}[f; \bar{F}]^2 \{\bar{f}(m+1) + \bar{f}(m)\}^{-1},$$

*where $\eta_z = 0$, $G(m) = |U|/m$, $\bar{F}$ is defined in (4.4) with $\Lambda \setminus \{z\}$ replaced by $U$ and $T_{xy}\eta$ is defined in (2.9). Here the constant is independent of $m/|U|$.*

Applying Lemma 4.2 with $U \to \Lambda\backslash\{z\}, \omega_{U,m} = \nu^{(z)}_{\Lambda,n}|_{\eta_z = n-m+1}$ and $\bar{f}(m) = E^{\nu^{(z)}_{\Lambda,n}}[f|\eta_z = n - m + 1] = \bar{f}_z(\eta_z = n - m + 1)$, we have proved Lemma 4.1. Finally, we prove Lemma 4.2.

**Proof of Lemma 4.2.** Recall $F(\eta_y)$ is defined in (4.4). From the definition of covariance,

$$\bar{f}(m) = E^{\omega_{U,m}}[f] = E^{\omega_{U,m}}[\bar{F}]^{-1}\left\{ - E^{\omega_{U,m}}[f; \bar{F}] + E^{\omega_{U,m}}[f\bar{F}]\right\}$$

For each $x \in \Lambda$ fixed, change the variable to $\eta = T_{x0}\xi$ where $T_{x0}$ is defined in (2.9). Hence

$$E^{\omega_{U,m}}[f\bar{F}] = \tilde{G}(m) Av_{x\in\Lambda} E^{\omega_{U,m+1}}[f(T_{x0}\xi)\xi_x],$$

where is some function of $m$ independent of $f$. Combining these two identities one has

$$\bar{f}(m) = -E^{\omega_{U,m}}[\bar{F}]^{-1} E^{\omega_{U,m}}[f;\bar{F}] + G(m+1) Av_{x\in\Lambda} E^{\omega_{U,m+1}}[f(T_{x0}\xi)\xi_x].$$

For some constant $G(m+1)$. The constant $G(m+1)$ can be determined by putting $f = 1$, namely $G(m+1) = L^d/m+1$. Define $E^{\chi_{m+1}}[t]$ for a family of functions $t = (t_x)_{x\in\Lambda}$ depending on $x$ by

$$E^{\chi_{m+1}}[t] = G(m+1) Av_{x\in\Lambda} E^{\omega_{U,m+1}}[t_x(\eta)\eta_x].$$

In particular, if $t$ is independent of $x$ then $E^{\chi_{m+1}}[t] = E^{\omega_{U,m+1}}[t]$. Hence $E^{\chi_{m+1}}[f(\eta)] = \bar{f}(m+1)$. Using these identities,

$$\bar{f}(m+1) - \bar{f}(m) = -E^{\omega_{U,m}}[\bar{F}]^{-1} E^{\omega_{U,m}}[f;\bar{F}] - E^{\chi_{m+1}}[(f(T_{*0}\eta) - f(\eta))]. \qquad (4.10)$$

*Step 2.* Let $f(T_{*0}\eta) - f(\eta)$ denote the family of functions $(f(T_{x0}\eta) - f(\eta))_{x\in\Lambda}$. From the elementary inequality

$$[\sqrt{\bar{f}(m+1)} - \sqrt{\bar{f}(m)}]^2 \le [\bar{f}(m+1) - \bar{f}(m)]^2/[\bar{f}(m+1) + \bar{f}(m)],$$

one has from (4.10),

$$[\sqrt{\bar{f}(m+1)} - \sqrt{\bar{f}(m)}]^2 \le A_3 + A_4$$

where

$$A_3 = 2\left\{E^{\chi_{m+1}}[f(T_{*0}\eta) - f(\eta)]\right\}^2 / [\bar{f}(m+1) + \bar{f}(m)]$$

$$A_4 = 2E^{\omega_{U,m}}[f;\bar{F}]^2 E^{\omega_{U,m}}[\bar{F}]^{-2} [\bar{f}(m+1) + \bar{f}(m)]^{-1}.$$

From the Schwarz inequality

$$\begin{aligned}
&\left\{E^{\chi_{m+1}}[f(T_{*0}\eta) - f(\eta)]\right\}^2 \\
&\le E^{\chi_{m+1}}\left[\left\{\sqrt{\bar{f}(T_{*0}\eta)} - \sqrt{\bar{f}(m)}\right\}^2\right] E^{\chi_{m+1}}\left[\left\{\sqrt{f(T_{*0}\eta)} + \sqrt{f(\eta)}\right\}^2\right] \\
&\le 2E^{\chi_{m+1}}\left[\left\{\sqrt{\bar{f}(T_{*0}\eta)} - \sqrt{\bar{f}(m)}\right\}^2\right] E^{\chi_{m+1}}\left[f(T_{*0}\eta) + f(\eta)\right]
\end{aligned}$$

By definition, $E^{\chi_{m+1}}[f(\eta)] = \bar{f}(m+1)$. Also, from (4.10), one has

$$E^{\chi_{m+1}}[f(T_{*0}\eta)] = E^{\omega_{U,m}}[\bar{F}]^{-1} E^{\omega_{U,m}}[f\bar{F}] + \bar{f}(m).$$

By definition, $\bar{F}$ is bounded above and $E^{\omega_{U,m}}[\bar{F}]^{-1}$ is bounded above by

$$E^{\omega_{U,m}}[\bar{F}]^{-1} \leq \text{ const.} E^{\omega_{U,m}}[Av_x(1-\eta_x)]^{-1} \leq \text{ const.}(1-m/L^d)^{-1}.$$

Together with the assumption $m/L^d \leq R/2$,

$$E^{\chi_{m+1}}[f(T_{*0}\eta)] \leq \text{ const.} \bar{f}(m)(1-m/L^d)$$

Hence

$$E^{\chi_{m+1}}\Big[f(T_{*0}\eta) + f(\eta)\Big][\bar{f}(m+1) + \bar{f}(m)]^{-1} \leq \text{ const.}(1-m/L^d)^{-1}$$

Hence we have proved $A_3 \leq A_1$ with $A_1$ defined in Lemma 4.2. Using again the bound on $E^{\omega_{U,m}}[\bar{F}]^{-1}$, we have $A_4 \leq A_2$. This proves Lemma 4.2.

We now return to the proof of Theorem 3.1.

*Step 3.* The exchange term in (4.6) can be bounded by the Dirichlet form using the following lemma from [SY]. We shall reproduce its proof in the Appendix.

**Lemma 4.3** *For every function $u$ on $\{0,1\}^{|\Lambda|}$ we have*

$$E^{\mu_{\Lambda,n}}[(T_{zy}u)^2] \leq \quad c_0|z-y| D_{\gamma_{zy}}(u),$$

*where $c_0$ is a constant depending on the Hamiltonian and $D_\gamma$ is defined in (3.4).*

Using Lemma 4.3, we can bound the exchange term in (4.6). Hence from (4.9) the entropy is bounded by

$$s(\bar{f}_z/\bar{\mu}^{(z)}_{\Lambda,n}) \leq \text{const.} \left\{ C_1|\log\rho| L^{1-d} \sum_{y\in\Lambda} D^{\gamma_{zy}}(\sqrt{f}) + \rho|\log\rho|\Phi \right\} \tag{4.11}$$

The first term on the right of (4.11) is, up to a factor $|\log\rho|$, of the form we need . Let us focus on the case where $\rho$ is bounded away from zero first.

Our next task is to bound the variance term $\Phi$. The following result, to be proved in the next section, is the key estimate of this paper.

**Theorem 4.4** *Suppose there is an $u(L)$ such that for any probability density $f$*

$$E^{\mu_{\Lambda_K,n}}[f\log f] \le u(L)L^2 Av_{b\in\Lambda_K} D_b(\sqrt{f}) \tag{4.12}$$

*for all $K \le L$. Then for any local function $h$ at $z \in \Lambda$ (with $\Lambda$ a cube of size $L\times 2L$ described before (3.1)), $\delta > 0$ and a large integer $\ell \ll L$, there is a constant $C_1$ such that*

$$\begin{aligned}\langle f; h\rangle^2_{\mu_{\Lambda,n}} \le\ & const.u(L) \sum_{|z-b|\le L/\ell} \Psi(|z-b|)D_b(\sqrt{f}) \\ & + \delta u(2L)L^{2-d}D_\Lambda(\sqrt{f}) + C_1 L^{2-d} D_\Lambda(\sqrt{f}).\end{aligned} \tag{4.13}$$

*Here $\Psi$ is defined in (3.5)*

We now apply Theorem 4.4 to the measure $\nu^{(z)}_{\Lambda,n}|_{\eta_z=0}$. The Dirichlet form on the right side of (4.13) are w.r.t this measure $\nu^{(z)}_{\Lambda,n}|_{\eta_z=0}$. Repeating the argument before (4.9), we have from the assumption $\rho \le 1/2$,

$$E^{\nu^{(z)}_{\Lambda,n}}[(T_b\sqrt{f})^2 \mid \eta_z = \alpha] \le \text{const.}\rho^{-1}E^{\mu_{\Lambda,n}}[(T_b\sqrt{f})^2], \qquad \alpha = 0,1.$$

We can now bound the middle term of $\Phi$ in (4.7) by

$$\begin{aligned}& E^{\nu^{(z)}_{\Lambda,n}}[f; g_z \mid \eta_z = 0]^2 \\ & \le \text{const.}E^{\nu^{(z)}_{\Lambda,n}}[f \mid \eta_z = 0]\Big[\rho^{-1}u(2L) \sum_{|z-b|\le L/\ell} \Psi(|z-b|)D_b(\sqrt{f}) \\ & \qquad + \delta\rho^{-1}u(2L)L^{2-d}D_\Lambda(\sqrt{f}) + C_1\rho^{-1}L^{2-d}D_\Lambda(\sqrt{f})\Big]\end{aligned}$$

The factor $E^{\nu^{(z)}_{\Lambda,n}}[f \mid \eta_z = 0] = \bar f_z(0)$ appears because $f$ is not necessary a probability density with respect to $\nu^{(z)}_{\Lambda,n}|_{\eta_z=0}$. Clearly, $\bar f_z(0)/[\bar f_z(1) + \bar f_z(0)] \le 1$. This gives a bound on the middle term of $\Phi$. Hence

$$\begin{aligned}& [\bar f_z(1) + \bar f_z(0)]^{-1} E^{\nu^{(z)}_{\Lambda,n}}[f; g_z \mid \eta_z = 0]^2 \\ & \le \text{const.}E^{\nu^{(z)}_{\Lambda,n}}[f \mid \eta_z = 0]\Big[\rho^{-1}u(2L) \sum_{|z-b|\le L/\ell} \Psi(|z-b|)D_b(\sqrt{f}) \\ & \qquad + \delta\rho^{-1}u(2L)L^{2-d}D_\Lambda(\sqrt{f}) + C_1\rho^{-1}L^{2-d}D_\Lambda(\sqrt{f})\Big]\end{aligned}$$

The last term in $\Phi$ can be bounded in the same way. For the first term in $\Phi$, from the Schwarz inequality, $E^{\nu^{(z)}_{\Lambda,n}}[f;\bar{F} \mid \eta_z = 1]^2 \leq Av_{x\in\setminus\{z\}}E^{\nu^{(z)}_{\Lambda,n}}[f;F_x \mid \eta_z = 1]^2$. We can now apply (4.13) to bound each term $E^{\nu^{(z)}_{\Lambda,n}}[f;F_x \mid \eta_z = 1]^2$. Repeating previous argument and averaging over $x \in \Lambda \setminus \{z\}$, we have

$$
\begin{aligned}
&[\bar{f}_z(1) + \bar{f}_z(0)]^{-1} E^{\nu^{(z)}_{\Lambda,n}}[f;\bar{F} \mid \eta_z = 1]^2 \\
&\quad \leq \text{const.}\rho^{-1}u(2L)Av_{x\in\Lambda} \sum_{|x-b|\leq L/\ell} \Psi(|z-b|)D_b(\sqrt{f}) \\
&\qquad + \delta\rho^{-1}u(2L)L^{2-d}D_\Lambda(\sqrt{f}) + C_1\rho^{-1}L^{2-d}D_\Lambda(\sqrt{f})
\end{aligned}
$$

The summation over $x$ and $b$ can estimated as in the step 2 of the proof of Theorem 2.1 in Sect 3. By choosing $\ell$ large enough, the first term on th eright side is bounded by $\delta\rho^{-1}u(2L)L^{2-d}D_\Lambda(\sqrt{f})$. Summarized, we have proved

$$
\Phi \leq \text{const.}\rho^{-1}Q_L
$$

where

$$
\begin{aligned}
Q_L &= C_1 L Av_{y\in\Lambda} D^{\gamma_{zy}}(\sqrt{f}) + \delta u(2L)L^{2-d}D_\Lambda(\sqrt{f}) + C_1 L^{2-d}D_\Lambda(\sqrt{f}) \\
&\quad + \text{const.}u(2L)\sum_{|z-b|\leq L/\ell} \Psi(|z-b|)D_b(\sqrt{f}).
\end{aligned}
\tag{4.15}
$$

Together with (4.11) we have thus proved Theorem 3.1 unless $\rho \to 0$. More precisely, we have

$$
s(\bar{f}_z/\bar{\mu}^{(z)}_{\Lambda,n}) = E^{\mu_{\Lambda,n}}[\bar{f}_z \log \bar{f}_z] \leq \rho|\log\rho|\{\sqrt{\bar{f}_z(0)} - \sqrt{\bar{f}_z(1)}\}^2 \leq |\log\rho| Q_L \tag{4.16}
$$

*Step 4.* Finally, we have to consider the low density region. For any integer $K \leq L$, let $\mathcal{F}_K$ be the $\sigma$-algebra generated by $\{\eta_y : y \notin \Lambda_K\}$ where $U_K$ is the cube in $\Lambda_L$ containing $z$ as a boundary point such that $dist(z, \Lambda_L \setminus U_K) = K+1$ and $z$ is a site in $\Lambda$ as defined in Theorem 3.1. Define the marginal density

$$
\bar{f}_{z,K}(\eta_z, \eta_y : y \notin \Lambda_K) = E^{\mu_{\Lambda,n}}[f \mid \eta_z, \mathcal{F}_K]
$$

Let $Q_K(\mathcal{F}_{\mathcal{K}})$ be the corresponding $Q$ in a cube of size $K$ with the boundary condition on $\Lambda\backslash\Lambda_K$ given by $\mathcal{F}_K$.

From the assumption (4.16) and the elementary bound $a^2 \le 2b^2 + 2(a-b)^2$,

$$\begin{aligned}\bar{f}_{z,K}(1) &= E^{\mu_{\Lambda,n}}[f\eta_z \mid \mathcal{F}_K]/E^{\mu_{\Lambda,n}}[\eta_z \mid \mathcal{F}_K] \\ &\le 2E^{\mu_{\Lambda,n}}[(1-\eta_z)f \mid \mathcal{F}_K]/E^{\mu_{\Lambda,n}}[(1-\eta_z) \mid \mathcal{F}_K] + 2Q_K(\mathcal{F}_{\mathcal{K}})/E^{\mu_{\Lambda,n}}[\eta_z \mid \mathcal{F}_K]\ ,\end{aligned} \tag{4.17}$$

Let $\rho_K$ be the density in the cube $U_K$ and denote $K^d\rho_K = n_K$. By definition,

$$E^{\mu_{\Lambda,n}}[\eta_z \mid \mathcal{F}_K] = E^{\mu_{U_K,n_K}}[\eta_z].$$

Applying (4.7) to our setting, we have

$$\text{const.} \le E^{\mu_{\Lambda,n}}[\eta_z \mid \mathcal{F}_K]/\rho_K \le \text{const.}.$$

Hence

$$E^{\mu_{\Lambda,n}}[f\eta_z \mid \mathcal{F}_K] \le \text{const.}[\rho_K/(1-\rho_K)]E^{\mu_{\Lambda,n}}[(1-\eta_z)f \mid \mathcal{F}_K] + 2Q_K(\mathcal{F}_{\mathcal{K}})\ , \tag{4.20}$$

Suppose

$$L^d/K^d \le \rho^{-1}/2. \tag{4.21}$$

Since $\rho_K \le \rho L^d/K^d$, we have $(1-\rho_K)^{-1} \le 2$. By definition, $E^{\mu_{\Lambda,n}}[Q_K(\mathcal{F}_{\mathcal{K}})] \le \text{const.}(K/L)Q_L$. Hence by taking the expectation of (4.20),

$$E^{\mu_{\Lambda,n}}[f\eta_z] \le \text{const.}\rho(L^d/K^d)E^{\mu_{\Lambda,n}}[f(1-\eta_z)] + 2(K/L)Q_L,$$

In other words, from the definition of $\bar{f}_z(1)$,

$$\bar{f}_z(1) \le 2(L^d/K^d)\bar{f}_z(0) + \rho^{-1}(K/L)Q_L = 2\bar{f}_z(0)\big[(L/K)^d + (K/L)(\frac{Q_L}{\bar{f}_z(0)\rho})\big] \tag{4.22}$$

provided (4.21) holds. Optimizing over $K$, we obtain that the optimizer is determined by

$$L/K = (\frac{Q_L}{\bar{f}_z(0)\rho})^{1/(d+1)}$$

Therefore,

$$\bar f_z(1) \le \text{const.}\bar f_z(0)\big(\frac{Q_L}{\bar f_z(0)\rho}\big)^{d/(d+1)} = \text{const.}\bar f_z(0)^{1/(d+1)}\big(\frac{Q_L}{\rho}\big)^{d/(d+1)} \tag{4.23}$$

provided that one chooses

$$1 \le L/K = \big(\frac{Q_L}{\bar f_z(0)\rho}\big)^{1/(d+1)} \le \rho^{-1/d}/2. \tag{4.24}$$

*Case 1*: $\bar f_z(0)\rho \le Q_L \le \bar f_z(0)\rho^{-1/d}/2^{d+1}$. Then (4.23) holds. Since $\rho \le 1/2$ and $\bar f_z(0)(1-\rho) \le 1$, one has $\bar f_z(0) \le (1-\rho)^{-1} \le 2$. Together with the bound $x \log x \le \text{const.}x^q$ if $x \ge 1$ and $q > 1$, one has

$$\rho \bar f_z(1) \log \bar f_z(1) \le \text{const.}\rho \bar f_z(1)^{(d+1)/d} \le \text{const.}Q_L.$$

provided $\bar f_z(1) \ge 1$. Since the left side is negative if $\bar f_z(1) < 1$, we do not need the assumption $\bar f_z(1) \ge 1$. Using the bound $\bar f_z(0)(1-\rho) \le 1$,

$$(1-\rho)\bar f_z(0) \log \bar f_z(0) \le -(1-\rho)\bar f_z(0)\log(1-\rho) \le 2\bar f_z(0)\rho \le 2Q_L.$$

Therefore, the entropy is bounded by

$$s(\bar f_z/\bar\mu^{(z)}_{\Lambda,n}) = \rho \bar f_z(1)\log \bar f_z(1) + \bar f_z(0)(1-\rho)\log \bar f_z(0) \le \text{const.}Q_L. \tag{4.25}$$

*Case 2*: $Q_L \le \bar f_z(0)\rho$. One has from (4.17) (with $K = L$ ), $\bar f_z(1) \le 4\bar f_z(0) \le 8$. From (4.3), the logarithmic factor in (4.16) can be omitted. This proves (4.25) and hence Theorem 3.1 in this case.

*Case 3*: $Q_L \ge \bar f_z(0)\rho^{-1/d}/2^{d+1}$. Hence $\bar f_z(0) \le 2^{d+1}\rho^{1/d}Q_L$ and we can replace (4.22) by

$$\bar f_z(1) \le 2^{d+2}\big[(L^d/K^d)\rho^{1/d} + \rho^{-1}(K/L)\big]Q_L.$$

Again, from (4.3) we can assume that $\bar f_z(1) \ge 4\bar f_z(0)$, for otherwise (4.25) follows immediately. Hence the entropy is bounded by $s \le 2\rho|\log\rho|\bar f_z(1)$ from (4.16). Let $L/K = \rho^{-1/d}/2$

so that (4.21) holds. Then $\bar{f}_z(1) \leq 16 \times 2^d \rho^{-(d-1)/d} Q_L$. Recall $s \leq 2\rho|\log \rho|\bar{f}_z(1)$. This proves (4.25) and concludes Theorem 3.1.

## V Large Deviation Estimates

We shall prove some large deviation estimates in this section. They are estimates based on the local limit theorem, which will be stated in Sect.VIII and proved in Sect. IX. These estimates will be useful in Sect.VI and Sect.VII. Because the technical nature of this section, we suggest the reader to skip this section until its results are needed in Sect.VI and Sect.VII. The key result is the Theorem 5.6 stated at the end of this section.

Suppose $\Lambda = \Lambda_L$ is a cube of width $L$ and $\mu_{L,n}$ is the canonical Gibbs state with number of particle $n$ and a fixed boundary condition. For applications in the next two sections, $\Lambda$ may be a rectangle as described in the paragraph before (3.1). All our results hold in that case with only notational changes. For notational simplicity we shall assume $\Lambda$ is cube of width $L$ for the rest of this section. *All results in sect. V-VII depend on the assumptions (A1-A4) unless otherwise stated.*

Let $U$ be a subcube of $\Lambda$ with width $L/\gamma$ for some constant $\gamma$ independent of $L$. We require that

$$\Lambda_L \setminus U = \text{a union of cubes of size } L/\gamma' \tag{5.1}$$

for some constant $\gamma'$ independent of $L$. In other words, we require that $\Lambda_L \setminus U$ has no "thin" region. This is because the mixing assumption can be violated if such pathological regions are allowed [MO]. Denote the density in $U$ by $\bar{\eta}_U$ and the expected density by

$$\rho_U^c = E^{\mu_{L,n}}[\bar{\eta}_U], \ \rho_U^{gc} = E^{\mu_{L,\lambda}}[\bar{\eta}_U] \tag{5.2}$$

If $U = \Lambda$, we shall drop the subscript $U$ and choose $\lambda$ such that $\rho^c = \rho^{gc} = \rho = n/L^d$.

**Theorem 5.1.** *Suppose $g$ is a smooth function on $(0,1)$ with*

$$g''(y) \leq c, \ g' \leq c, \ g(\rho) = 0 \ . \tag{5.3}$$

*Let*

$$\zeta(y) = g(y) - g'(\rho_U^c)(y - \rho_U^c) - g(\rho_U^c)$$

*Then*

$$\beta^{-1} L^{-d} \log E^{\mu_{\Lambda,n}} \Big\{ \exp\big[\beta L^d \{\zeta(\bar{\eta}_U)\}\big] \Big\} \leq \delta\beta \tag{5.4}$$

*provided that*

$$\beta \leq \beta_0 \ll 1. \tag{5.5}$$

*Here $\delta$ is a small constant and $\beta_0$ is a fixed constant. Furthermore, for some constant $C$*

$$\langle f; g(\bar{\eta}_U)\rangle^2 \leq C s(f) \tag{5.6}$$

*for any probability density $f$ with respect to $\mu_{L,n}$.*

Almost all constants for the rest of this paper depend on $\gamma$. Some times we obtain explicitly dependence on $\gamma$. All our results hold with little changes regardless of the value of $\gamma$ provided $\gamma$ is independent of $L$, which is the assumption for the rest of this paper.

We need the following lemmas to prove Theorem 5.1. The following Lemma 5.2 is useful when cutoffs are needed. For the rest of this section we shall use $\epsilon$ or $\delta$ to denote small positive numbers and use $\beta$ to denote arbitrary positive number.

**Lemma 5.2** *Let $\nu$ be a probability measure. Suppose $g$ is a function satisfying*

$$\parallel g - v \parallel_\infty \leq L^{-(d/2)-\epsilon} \tag{5.7}$$

*where $v$ denotes the expectation of $g$. Then for $\beta \leq \beta_0$ with $\beta_0$ independent of $L$*

$$\beta^{-1} L^{-d} \log E^{\nu} \big\{ \exp\big[\beta L^d (g - v)\big] \big\} \leq \delta\beta \tag{5.8}$$

*and $\delta \to 0$ as $L \to \infty$. If $g$ only satisfies (5.7) with $\epsilon = 0$ then (5.8) still holds for some $\delta$ independent of $L$.*

**Proof of Lemma 5.2.** *Case 1*: $\beta \geq L^{-(d+\epsilon)/2}$. From the assumption on $g$,

$$\beta^{-1} L^{-d} \log E^{\nu} \big\{ \exp\big[\beta L^d (g - v)\big] \big\} \leq \parallel g - v \parallel_\infty \leq L^{-(d/2)-\epsilon} \leq \delta\beta.$$

*Case 2*: $\beta \le L^{-(d+\epsilon)/2}$. Hence $\beta L^d \parallel g - v \parallel_\infty \le L^{-\epsilon/2}$. We can expand the exponential up to the second order to have

$$\beta^{-1} L^{-d} \log E^\nu \{ \exp[\beta L^d (g - v)] \} \le \beta L^d E^\nu [(g - v)^2] \le \beta L^{-\epsilon}$$

The higher order order terms are even smaller. This concludes Lemma 5.2. □

Recall the entropy bound

$$\int f X d\mu \le \beta^{-1} L^{-d} \log \int \exp[\beta L^d X] d\mu + \beta^{-1} s(f) \tag{5.9}$$

for any probability density $f$. This bound is a simple consequence of Jensen's inequality. From Lemma 5.2 and (5.9), a function $X$ with small $L^\infty$ norm satisfying (5.7) can be bounded by

$$\langle f; X \rangle \le \delta\beta + \beta^{-1} s(f) \tag{5.10}$$

Optimizing $\beta$, we have

$$\langle f; X \rangle^2 \le \delta s(f)$$

This will be sufficient to bound $\langle f; X \rangle^2$ for all purposes in this paper. Hence any term satisfying (5.7) is negligible and will be dropped for the rest of this paper.

**Proof of Theorem 5.1, Part I** We first prove (5.4). Suppose $\rho \le L^{-d/2+\epsilon}$. Then

$$\bar\eta_U \le \gamma^d \rho \le \gamma^d L^{-(d/2)+\epsilon} .$$

From (5.3) one has

$$\zeta(y) \le \text{const.} |y - \rho_U^c|^2 := t(y) \tag{5.11}$$

It follows that $|\zeta(\bar\eta_U)| \le L^{-d+2\epsilon}$ and (5.7) is satisfied. Hence (5.4) follows from Lemma 5.2. *For the rest of the proof in Part I, we shall assume that* $\rho \ge L^{-d/2+\epsilon}$. The following Lemma is the key input.

**Lemma 5.3** *Suppose $\mu_{L,n}$ is the canonical Gibbs state with density $\rho = n/L^d$ satisfying $\rho \ge L^{-d/2}$. Recall $\rho_U^c = E^{\mu_{L,n}}[\bar\eta_U]$. Then there is a large constant $C > 0$ such that*

$$P^{\mu_{L,n}} \left[ |\bar\eta_U - \rho_U^c| \ge |y - \rho_U^c| \right] \le \exp[-const. (L^d/\gamma^d) |y - \rho_U^c|^2]$$

*provided that* $|y-\rho_U^c| \geq C\gamma^{d/2}L^{-(d/2)}$.

Assuming this lemma, we now continue the proof of Theorem 5.1.

Suppose first $\beta \leq L^{-10d}$. We can expand the exponential in (5.4) and prove Theorem 5.1 directly. This is straightforward and we omit the detail.

Suppose $L^{-10d} \leq \beta \leq \beta_0$. First, we consider the region $|\bar\eta_U - \rho_U^c| \geq L^{-(d/2)+\epsilon}$. From Lemma 5.3 and (5.11)

$$\begin{aligned} & E^{\mu_{L,n}}\Big\{ \exp\big[\beta L^d \zeta(\bar\eta_U)\big] \mathbf{1}(|\bar\eta_U - \rho_U^c| \geq L^{-(d/2)+\epsilon})\Big\} \\ \leq & \sum_{y=j/L^d, j\in\mathbb{Z} : (y-\rho_U^c)\geq L^{-(d/2)+\epsilon}} \exp\big[\beta L^d t(y)\big] P^{\mu_{L,n}}(\bar\eta_U = y) \\ & + \sum_{y:(y-\rho_U^c)\leq -L^{-(d/2)+\epsilon}} \exp\big[\beta L^d t(y)\big] P^{\mu_{L,n}}(\bar\eta_U = y) \end{aligned} \tag{5.12}$$

We can overestimate $P^{\mu_{L,n}}(\bar\eta_U = y)$ by $P^{\mu_{L,n}}(\bar\eta_U \geq y)$. From Lemma 5.3 and (5.11), we have

$$\exp\big[\beta L^d t(y)\big] P(\bar\eta_U \geq y) \leq \exp[-\text{const}.L^\epsilon].$$

for $(y-\rho_U^c) \geq L^{-(d/2)+\epsilon}$. Hence

$$E^{\mu_{L,n}}\Big\{ \exp\big[\beta L^d \zeta(\bar\eta_U)\big] \mathbf{1}(|\bar\eta_U - \rho_U^c| \geq L^{-(d/2)+\epsilon})\Big\} \leq \exp[-\text{const}.L^\epsilon].$$

This proves that the contribution of this region to (5.4) is negligible.

Finally we have to estimate the contribution of the region $|\bar\eta_U - \rho_U^c| \leq L^{-(d/2)+\epsilon}$. From (5.11), we have that $|\zeta(\bar\eta_U)| \leq L^{-d+2\epsilon}$ for some small positive constant $\epsilon$ . Hence $\zeta$ satisfies (5.7) and (5.4) follows from Lemma 5.2. □

We now prove Lemma 5.3. First we introduce some notations. Define the pressure

$$\begin{aligned} \phi(\theta,\rho) &= (\gamma^d/L^d)\log E^{\mu_{L,n}}\left[\exp(\theta\sum_{x\in U}\eta_x)\right] = (\gamma^d/L^d)\log M(\theta,\rho), \\ M(\theta,\rho) &= E^{\mu_{L,n}}\left[\exp\left(\theta\sum_{x\in U}\eta_x\right)\right] , \end{aligned} \tag{5.13}$$

where $\rho = \bar{\eta} = n/L^d$ denotes the total density in $\Lambda_L$. From the Chebyshev inequality, we have

$$(L^d/\gamma^d) \log P^{\mu_{L,n}}[\bar{\eta}_U \geq y] \leq -h(y, \rho) := -\sup_\theta \{\theta y - \phi(\theta, \rho)\} \tag{5.14}$$

Similarly, define

$$\begin{aligned} R(\theta, \lambda) &= (\gamma^d/L^d) \log E^{\mu_{L,\lambda=0}} \left[ \exp\left( \lambda \sum_{x \in \Lambda} \eta_x + \theta \sum_{x \in U} \eta_x \right) \right] , \\ p(\theta, \rho) &= -\sup_\lambda \left[ \lambda \gamma^d \rho - R(\theta, \lambda) \right] . \end{aligned} \tag{5.15}$$

The following Lemma gives a relation between $\phi$ and $p$.

**Lemma 5.4** *If $\rho \geq L^{-d/2}$ then*

$$| \phi(\theta, \rho) - (p(\theta, \rho) - p(\theta = 0, \rho)) | \leq const.(\gamma^d/L^d).$$

Lemma 5.4 will be proved in section 8. We need some more notations. Recall that $\mu_{L,\lambda}$ is the Gibbs state on a cube $\Lambda_L$ with chemical potential $\lambda$ and some fixed boundary condition. Recall $U$ is a subcube of size $L/\gamma$ with the property that $\Lambda_L \setminus U$ is "fat" (5.1). Let $\mu_{L,\theta,\lambda}$ be the probability measure with density relative to $\mu_{L,\lambda=0}$ given by

$$\exp(\theta \sum_{x \in U} \eta_x + \lambda \sum_{x \in \Lambda_L} \eta_x )\, Z(\theta, \lambda)^{-1} \tag{5.16}$$

where $Z(\theta, \lambda)$ is the normalization. Clearly, when $\theta = 0$ the measure $\mu_{L,\theta=0,\lambda}$ reduces to $\mu_{L,\lambda}$.

We claim that $\mu_{L,\theta,\lambda}$ satisfies the mixing condition (A1) for all $\theta$ and $\lambda$ and for all cubes if $\mu$ satisfies the mixing condition (A1) for all $\lambda$ and for all cubes. The issue is standard, therefore, we only sketch the idea. Roughly speaking, our goal is to prove the correlation function of two local functions $f_x$ and $g_y$ at $x$ and $y$ resp. decay exponentially with $|x - y|$. Suppose $x \in U$ and $y \in \Lambda_L \setminus U$. Then

$$E^{\mu_{L,\theta,\lambda}}[f_x : g_y] = E^{\mu_{L,\theta,\lambda}} \left[ g_y : E^{\mu_{L,\theta,\lambda}}[f_x | \eta_x, x \in \Lambda_L \setminus U] \right]$$

The conditional expectation, $E^{\mu_{L,\theta,\lambda}}[f_x|\eta_x, x \in \Lambda_L \setminus U]$ depends on the configurations on $\Lambda_L \setminus U$ only through the boundary condition. From (A.1) the dependence of $E^{\mu_{L,\theta,\lambda}}[f_x|\eta_x, x \in \Lambda_L \setminus U]$ on the configuration at a fixed site $z$ in the boundary is of order $\exp[-\text{const.}|z-x|]$. This proves

$$\left| E^{\mu_{L,\theta,\lambda}}[f_x : g_y] \right| \le \exp\left[-C|x-y|\right]$$

The other cases, $x, y \in U$ or $x, y \in \Lambda_L \setminus U$, can be proved in a similar way.

The following Lemma provides a bound on a special correlation needed in the proof of Lemma 5.3.

**Lemma 5.5** *If* $\rho \ge L^{-d/2}$ *then*

$$(L^d/\gamma^d)\langle \bar{\eta}_U; \bar{\eta}_U \rangle_{\mu_{\theta,\lambda}} \le C$$

Lemma 5.5 will also be proved in section 8. Return to the proof of Lemma 5.3.

**Proof of Lemma 5.3** We first bound the probability of the event in Lemma 5.3 by (5.14). From Lemma 5.4, up to an error $(\gamma^d/L^d)$, we can replace $h$ by $f - p(0,\rho)$, where $f$ is the Legendre transform of $p$, namely

$$f(y,\rho) := -\sup_\theta \{\theta y - p(\theta,\rho)\}$$

To prove Lemma 5.3, from (5.14) it suffices to prove that

$$f(y,\rho) - p(0,\rho) \ge C(y-\rho_U^c)^2$$

for $|y-\rho_U^c| \ge C\gamma^{d/2}L^{-(d/2)}$. The error $\gamma^d L^{-d}$, is negligible in this region since $\gamma^d L^{-d} < (L^d/\gamma^d)|y-\rho_U^c|^2$. We claim that

$$|\rho_U^{gc} - \rho_U^c| \le \gamma^d L^{-d/2} \tag{5.17}$$

Assuming this bound, we only have to prove that

$$f(y,\rho) - p(0,\rho) \ge C(y-\rho_U^{gc})^2 \tag{5.18}$$

From Lemma 5.4,

$$\mid \phi(\theta,\rho) - (p(\theta,\rho) - p(\theta=0,\rho)) \mid \leq \text{const.}(\gamma^d/L^d).$$

if $\rho \leq L^{-d/2}$. Note that from the convexity of $\phi$,

$$-\theta^{-1}[\phi(-\theta,\rho) - \phi(0,\rho)] \leq \langle \bar\eta_U \rangle_{\mu_{L,n}} \leq \theta^{-1}[\phi(\theta,\rho) - \phi(0,\rho)]$$

Similar bound holds if $\langle \bar\eta_U \rangle_{\mu_{L,n}} = \rho_U^c$ is replaced by $\langle \bar\eta_U \rangle_{\mu_{L,\lambda}} = \rho_U^{gc}$ and $\phi$ is replaced by $p$. Choosing $\theta = L^{-d/2}\gamma^{d/2}$, we obtain (5.17) provided $\rho \geq L^{-d/2}$. Clearly, (5.17) holds trivially if $\rho \leq L^{-d/2}$.

We now prove the (5.18). From the Taylor Theorem, it suffices to prove

$$\begin{aligned} &f(\rho_U^{gc},\rho) = p(0.\rho) \\ &\partial f(y,\rho)/\partial y|_{y=\rho_U^{gc}} = 0 \\ &\partial^2 f(y,\rho)/\partial y^2 \geq C \end{aligned}$$

The first two identities follow from the definition of $\rho_U^{gc}$. We now prove the last bound. By definition,

$$\partial^2 f(y,\rho)/\partial y^2 = \Big(\partial^2 p(\theta,\rho)/\partial\theta^2\Big)^{-1}$$

From simple calculation,

$$\begin{aligned} \partial p(\theta,\rho)/\partial\theta &= \langle \bar\eta_U \rangle_{\mu_{L,\theta,\lambda}} \\ \partial^2 p(\theta,\rho)/\partial\theta^2 &= (L^d/\gamma^d)\big[\langle \bar\eta_U;\bar\eta_U \rangle_{\mu_{L,\theta,\lambda}} - \langle \bar\eta;\bar\eta_U \rangle^2_{\mu_{L,\theta,\lambda}} \langle \bar\eta;\bar\eta \rangle^{-1}_{\mu_{L,\theta,\lambda}}\big] \\ &\leq (L^d/\gamma^d)\langle \bar\eta_U;\bar\eta_U \rangle_{\mu_{L,\theta,\lambda}} \leq C \end{aligned} \tag{5.19}$$

where the last bound follows from Lemma 5.5. This proves $\partial^2 f(y,\rho)/\partial y^2 \geq C$ and concludes Lemma 5.3. □

**Proof of Theorem 5.1, Part II** We now prove (5.6) assuming (5.4) holds. From the Schwarz inequality

$$\beta^{-1}L^{-d}\log E^{\mu_{L,n}}\Big\{\exp\big[\beta L^d\{g(\bar\eta_U) - g(\rho_U^c)\}\big]\Big\} \leq Q_1 Q_2 \tag{5.20}$$

where

$$Q_1 = \beta^{-1} L^{-d} \log E^{\mu_{L,n}} \Big\{ \exp\big[2\beta L^d \{\zeta(\bar\eta_U)\}\big] \Big\}^{1/2}$$

$$Q_2 = \beta^{-1} L^{-d} \log E^{\mu_{L,n}} \Big\{ \exp\big[2\beta L^d \{g'(\rho_U^c)(\bar\eta_U - \rho_U^c)\}\big] \Big\}^{1/2}$$

and $\zeta$ is defined Theorem 5.1. We can use (5.4) to bound $Q_1$ provided $\beta \le \beta_0$.

By definition of $\phi$,

$$Q_2 = \beta^{-1}\phi(\theta, \rho) - \theta\rho_U^c, \ \theta = 2\beta\gamma^d g'(\rho_U^c)$$

We shall prove

$$Q_2 \le C\beta \tag{5.21}$$

Suppose that $\beta \ge CL^{-d/2}$. Then $\beta^2 \le CL^{-d}$. Hence

$$\beta^{-1}\Big|\phi(\theta,\rho) - [p(\theta,\rho) - p(\theta = 0, \rho)]\Big| \le \beta^{-1}L^{-d} \le \text{const.}\beta$$

Therefore, it suffices to prove

$$\beta^{-1}\Big|[p(\theta,\rho) - p(0,\rho)] - \theta\rho_U^c\Big| \le \text{const.}\beta$$

From (5.17) and assumption on $\beta$, the error

$$|\partial p(\theta,\rho)/\partial\theta|_{\theta=0} - \rho_U^c| = |\rho_U^{gc} - \rho_U^c| \le CL^{-d/2}$$

is bounded by $C\beta$. We now expand $p$ in $\theta$. Since the second derivative of $p$ is bounded from Lemma 5.5, we have

$$\beta^{-1}\Big|\phi(\theta,\rho) - [p(\theta,\rho) - p(\theta = 0, \rho)]\Big| \le C|\rho_U^{gc} - \rho_U^c| + C\beta$$

We have thus proved (5.21) provided $\beta \ge CL^{-d/2}$.

Suppose $\beta \le CL^{-d/2}$. We can apply Lemma 5.2 in this case. More precisely, let $\beta = \tilde\beta L^{-d/2}$ and replace $g$ by $\tilde g = L^{-d/2} g$. Hence we can apply Lemma 5.2 with $\tilde\beta$ and $\tilde g$. This proves (5.21) in this case. Putting the bounds on $Q_1$ and $Q_2$ together, we have proved

$$\beta^{-1} L^{-d} \log E^{\mu_{L,n}} \Big\{ \exp\big[\beta L^d \{g(\bar\eta_U) - g(\rho_U^c)\}\big] \Big\} \le C\beta \tag{5.22}$$

provided $\beta \le \beta_0$. But (5.22) holds trivially if $\beta \ge \beta_0$. Hence the condition $\beta \le \beta_0$ can be dropped.

Let $X = g(\bar\eta_U) - g(\rho_U^c)$. From (5.10) and (5.22) one has

$$\langle f; X\rangle \le C\beta + \beta^{-1}s(f)$$

Let $\beta = C_1\sqrt{s}(f)$. We have

$$\langle f; X\rangle^2 \le \text{const}.s(f).$$

This concludes (5.6) and finishes the proof of Theorem 5.1. □

We now provide a class of functions satisfying the assumption (5.3). Suppose $g$ is a local function at $z \in \Lambda_L$ which may be near the boundary of $\Lambda_L$. Define $g_U$ by

$$g_U = E^{\mu_{L,n}}[g \mid \mathcal{F}_U]$$

where $\mathcal{F}_U$ is the $\sigma$-algebra generated by $\{\eta_y; y \notin U\}$. Since the total number of particles in $\mu_{L,n}$ is fixed, $\bar\eta_U = Av_{x\in U}\ \eta_x$ is measurable with respect to $\mathcal{F}_U$. By definition $g_U$ depends on the boundary condition on $\partial U$. Let $\omega = (\omega_1, \omega_2)$ denote the boundary condition of $U$ with $\omega_1$ ($\omega_2$ resp.) denoting the boundary condition on $\partial\Lambda_L \cap \partial U$ ($\partial U \setminus \partial\Lambda_L$ resp.) . We require that

$$dist(z, \partial U \setminus \partial\Lambda_L) = L(4\gamma)^{-1}. \tag{5.23}$$

This assumption will hold for all applications in this paper. Let $\hat g_U(y)$ be the expectation of $g$ with respect to the *grand* canonical Gibbs state with the boundary condition $\omega$ and density $y$, i.e.,

$$\hat g_U(y) = E^{\mu_{L,\lambda}}[g], \tag{5.24}$$

where $\lambda$ is chosen so that the density is $y$, i.e., $E^{\mu_{L,\lambda}}[\bar\eta_U = y]$. From (A.4)

$$|g_U - \hat g_U(\bar\eta_U)| \le O(L^{-d/2-\epsilon}). \tag{5.25}$$

From (5.23) and the mixing condition (A1) the dependence on the boundary condition $\omega_2$ is bounded by $O(L^{-d})$. Let

$$\tilde g_U(y) = E^{\mu_{L,\omega_1,\lambda}}[g], \tag{5.26}$$

where $\mu_{L,\omega_1,\lambda}$ is the measure on $U$ with boundary condition $\omega_2 = 0$ and $\lambda$ chosen such that $E^{\mu_{L,\omega_1,\lambda}}[\bar\eta_U] = y$. Then

$$|g_U - \tilde g_U(\bar\eta_U)| \le O(L^{-d/2-\epsilon}). \tag{5.27}$$

From (A2), $\tilde g_U$ satisfies (5.3). Furthermore, from the remark after (5.10) the difference between $g_U$ and $\tilde g_U$ is negligible. We have thus proved the following Theorem.

**Theorem 5.6** *With previous notations and assumptions, for any probability density $f$ with respect to $\mu_{L,n}$*

$$\langle f; g_U(\bar\eta_U)\rangle^2_{\mu_{L,n}} \le Cs(f) \tag{5.28}$$

*for some constant $C$. Furthermore,*

$$\beta^{-1}L^{-d}\log E^{\mu^c_{L,n}}\Big\{\exp\big[\beta L^d\{\tilde g_U(\bar\eta_U) - \tilde g'_U(\rho^c_U)(\bar\eta_U - \rho^c_U) - \tilde g_U(\rho^c_U)\}\big]\Big\} \le \delta\beta \tag{5.29}$$

*provided (5.5) holds.*

## VI Two-Block Estimates and Multiscale Analysis

We now prove Theorem 4.4. The basic ingredients for proving Theorem 4.4 are a multi-scale estimate, a large deviation bound (Theorem 5.6) and a precise statement of the two-block estimate. In this section, we shall use the two-block estimate and multiscale analysis to prove part of Theorem 4.4. We then use a large deviation estimate from sect. 5 to conclude Theorem 4.4 in sect. 7. Let us first define some notations.

Recall that $\Lambda$ is a cube of size $L_1 \times L_2$ with $L_i$ about the size of $L$ as defined before (3.1). Suppose q is a local function at $z \in \Lambda$ which may be near the boundary or corners of $\Lambda$. Fix an integer $\ell$. Let $q^{(j)}$ be defined by

$$g^{(j)} = E^{\mu_{L,n}}[g \mid \mathcal{F}_j] \tag{6.1}$$

where $\mathcal{F}_j$ is the $\sigma$-algebra generated by $\{\eta_y; y \notin \Lambda_j$ with $\Lambda_j = \{y \in \Lambda : |y-z| > \ell^j\}\}$. Since the total number of particles in $\mu_{L,n}$ is fixed, $\bar\eta_j = Av_{x\in\Lambda_j}\eta_x$ is measurable with respect to $\mathcal{F}_j$. Suppose $\ell^{m+2d+2} = 2L$ for some integer $m$. The choice of the strange exponent $m + 2d + 2$ is for convenience and will become clear later.

To avoid pathological cases, we have to redefine the cubes $\Lambda_j$ slightly. Since the site $z$ may be in the corner, the set $\Lambda_{j+1} \setminus \Lambda_j$ may not be a "fat" region according to (5.1) (with $\Lambda$ replaced by $\Lambda_j$ and $L/\gamma$ replaced by $\ell^j$). In this case, we shall enlarge $\Lambda_j$ to $\Lambda_{j+1}$ to eliminate the pathological case. Certainly, there are only finite number of pathological cases and the changes will not affect our estimates. We shall not comment on this further in this article.

For convenience of notation, we use

$$U = \Lambda^{(m)},\ \bar{\eta}_U = \bar{\eta}_m,\ g_U = g^{(m)}, \gamma = \ell^{2d+2}. \tag{6.2}$$

Also by (5.23), we can replace $g_U$ by $\tilde{g}_U$ whenever needed. We remind the reader that $\tilde{g}_U(y)$ is nothing but the expectation of $g$ with respect to the grand canonical Gibbs state on $U$ with $E^{\mu^{gc}_{L,\lambda}}[\bar{\eta}_U] = y$ and some fixed boundary condition described in the paragraph before Theorem 5.6.

**Lemma 6.1.** *Recall the definition of u(L) in (3.1). Assume as usual that the mixing condition (A1-A4) hold. Then*

$$\langle f; g\rangle^2_{\mu_{L,n}} \le const.\ell^{2d} u(L) \sum_{|b-z|\le L/\ell^{2d+2}} \Psi(|z-b|) D_b(\sqrt{f}) + 2\langle f; g_U\rangle^2_{\mu_{L,n}}$$

*where $g_U$ is defined in (6.2) with $\ell^{m+2d+2} = L$.*

**Proof** Recall the following martingale decomposition for any function $g$,

$$\langle f; g\rangle_{\mu_{L,n}} = \sum_{j=0}^{m-1} \langle f; g^{(j)} - g^{(j+1)}\rangle_{\mu_{L,n}} + \langle f; g_U\rangle_{\mu_{L,n}}\ , \tag{6.3}$$

and for each $j$ fixed,

$$\langle f; g^{(j)} - g^{(j+1)}\rangle_{\mu_{L,n}} = E^{\mu_{L,n}}[E^{\mu_{L,n}}[f(g^{(j)} - g^{(j+1)})|\mathcal{F}_{j+1}]. \tag{6.4}$$

Since the total number of particles in $\Lambda$ is fixed, the total number of particles in $\Lambda_j$ is also fixed once $\mathcal{F}_j$ is given. From Corollary 5.7 and the definition of $u$ (3.1),

$$\langle f; g^{(j)} - g^{(j+1)}\rangle^2_{\mu_{L,n}} \le \text{const.}\ell^{2d}u(L)\ell^{j(2-d)}D_{\Lambda_j}(\sqrt{f}). \tag{6.5}$$

Using this bound and the Schwarz inequality, for any positive $\beta$ one can bound the variance by

$$\begin{aligned}&\sum_{j=0}^{m-1}\langle f; g^{(j)} - g^{(j+1)}\rangle_{\mu_{L,n}} \le \text{const.}\beta\sum_{j=0}^{m-1}(m+2-j)^{-2}\\ &+\text{const.}\beta^{-1}\ell^{2d}u(L)\sum_{b:|b-z|\le L/\ell^{2d+2}}\ \sum_{j:\ell^j\ge|b-z|}(m+2-j)^2\ell^{j(2-d)}D_b(\sqrt{f})\ .\end{aligned} \tag{6.6}$$

The summation over $j$ can be bounded by

$$\sum_{j:\ell^j\ge|b-z|}(m+2-j)^2\ell^{j(2-d)} \le \Psi(|z-b|) \tag{6.7}$$

where $\Psi$ is defined in (3.5). Optimizing over $\beta$ one has

$$\Big\{\sum_{j=0}^{m-1}\langle f; g^{(j)} - g^{(j+1)}\rangle_{\mu_{L,n}}\Big\}^2 \le \text{const.}\ell^{2d}u(L)\sum_{b:|b-z|\le L/\ell^{2d+2}}\Psi(|z-b|)D_b(\sqrt{f})\ .$$

This concludes the proof of Lemma 6.1. □

From Lemma 6.1, it suffices to bound the last term $\langle f; g_U\rangle^2_{\mu_{L,n}}$ in order to conclude Theorem 4.4. We state it as the following Lemma.

**Lemma 6.2.** *With the same notations and assumptions as in the previous Lemma,*

$$\langle f; g_U\rangle^2_{\mu_{L,n}} \le \mathit{const.}[C(\delta)+\delta u(2L)]L^{2-d}D_\Lambda(\sqrt{f}) \tag{6.8}$$

*where $\delta$ is a small constant and $g_U$ is defined in (6.2) with $\ell^{m+2d+2}=2L$.*

In this section, we shall only prove parts of Lemma 6.2, stated as the following Lemma 6.3. The rest will be presented in the next section.

**Lemma 6.3** *Suppose that for some constant $\delta_1 > 0$,*

$$\big[u(2L)L^{2-d}D_\Lambda(\sqrt{f})\big] \ge \delta_1 \tag{6.9}$$

*Then*

$$\langle f; g_U(\bar{\eta}_U)\rangle^2_{\mu_{L,n}} \leq const. L^{2-d} D_\Lambda(\sqrt{f}). \tag{6.10}$$

*Here the constant depends on $\gamma$ and $\delta_1$.*

We now prove Lemma 6.3. The following Lemma 6.4 will be needed. For those familiar with hydrodynamic limits, it can be understood as a statement of the two–block estimate.

**Lemma 6.4** *Suppose $\xi = (\eta, \zeta)$ are configurations on $\Lambda_k \cup \Lambda_k$. Let $\nu$ be a canonical Gibbs measure on $\Lambda_k \cup \Lambda_k$ with a fixed boundary condition. Define $\xi^{(x)} = (\eta^{(x)}, \zeta^{(x)})$ by*

$$\eta^{(x)}|_y = \begin{cases} \eta_y, & \text{if } y \neq x \\ \zeta_x, & \text{if } y = x \end{cases}$$

*and*

$$\zeta^{(x)}|_y = \begin{cases} \zeta_y, & \text{if } y \neq x \\ \eta_x, & \text{if } y = x \end{cases}$$

*In other words, $\xi^{(x)}|$ denotes the configuration obtained by exchanging $\eta_x$ and $\zeta_x$. Define the Dirichlet form $D_{\eta,\zeta}$ by*

$$D_{\eta,\zeta}(f) = D^{(\eta)} + D^{(\zeta)} + \sum_{x \in \Lambda_k} D^{(x)}_{\eta,\zeta}(f) \tag{6.11}$$

*where $D^{(\eta)}$ ($D^{(\zeta)}$ resp.) is the usual Dirichlet form on the configuration $\eta$ ($\zeta$ resp.) alone and*

$$D^{(x)}_{\eta,\zeta}(f) = \sum_{x \in \Lambda_k} \int [f(\xi^{(x)}) - f(\xi)]^2 \, d\nu(\xi), \tag{6.12}$$

*Then there are constants $C_1(k)$ and $C_2(k)$ such that*

$$\int f(\xi)(\bar{\eta} - \bar{\zeta})^2 d\nu(\xi) - C_2(k) \leq C_1(k) D_{\eta,\zeta}(\sqrt{f}), \tag{6.13}$$

*and $\lim_{k\to\infty} C_2(k) = 0$.*

**Proof of Lemma 6.4** First, the constant $C_1(k)$ can be chosen as large as possible (depending on $k$). Let $C_2(k) = \int (\bar{\eta} - \bar{\zeta})^2 d\nu$ be the expectation of $(\bar{\eta} - \bar{\zeta})^2$ with respect to $\nu$. measure. Hence $\lim_{k\to\infty} C_2(k) = 0$. Note that the right side of (6.13) vanishes only if the left side vanishes. Since the configuration space is finite, there exists a constant $C_1(k)$ so that (6.13) holds. More precise dependence of $C_1(k)$ and $C_2(k)$ can be obtained. But we will not need these bounds here. □

**Proof of Lemma 6.3** For any two cubes $\alpha, \beta$ of size $k$ in $\Lambda$ one has from Lemma 6.4

$$\int f(\bar{\eta}_\alpha - \bar{\eta}_\beta)^2 d\mu_{L,n} \le C_1(k) D_{\alpha,\beta}(\sqrt{f}) + C_2(k)$$

Here $D_{\alpha,\beta}$ is defined in (6.11) with the cubes $\alpha, \beta$ taking the roles of $\Lambda_k \times \Lambda_k$ in the lemma. By definition $D_{\alpha,\beta}$ involves exchanging particles in the cubes $\alpha$ and $\beta$. From Lemma 4.3, we can bound these exchanges by the usual Dirichlet form with only nearest neighbor exchanges. Averaging over $\alpha$ in $\Lambda_m = U$ and $\beta$ in $\Lambda$ and then using the Schwarz's inequality, we have

$$\int f(\bar{\eta}_m - \rho)^2 d\mu_{L,n} \le C_1(k) L^{2-d} D_\Lambda(\sqrt{f}) + C_2(k) \tag{6.14}$$

From (5.27), we can replace $g_U$ by $\tilde{g}_U$ in proving Lemma 6.3. From (A2), $\tilde{g}$ is uniformly Lipschitz continuous. Therefore, one has from (6.14)

$$\begin{aligned}\langle f; \tilde{g}_U(\bar{\eta}_U)\rangle^2_{\mu_{L,n}} &\le E^f \left[\tilde{g}_U(\bar{\eta}_U) - \tilde{g}_U(\rho)\right]^2 \le \text{const.} E^f \left[(\bar{\eta}_U - \rho)^2\right] \\ &\le \text{const.}[C_1(k) L^{2-d} D_\Lambda(\sqrt{f}) + C_2(k)]\end{aligned}$$

Choosing $k$ large enough and using the assumption (6.9) we conclude lemma 6.3.

## VII Proof of Lemma 6.2

Let us summarize what we have proved so far. From Lemma 6.3 we can assume that $\beta$ satisfies

$$\left[u(2L)L^{2-d}D_\Lambda(\sqrt{f})\right] \le \delta_1 \tag{7.1}$$

Recall the definition of $\tilde{g}_U$ in (5.26). As remarked at the end of sect. 5, we can replace $g_U$ by $\tilde{g}_U$. Hence we have to prove

$$\langle f; \tilde{g}_U\rangle^2_{\mu_{L,n}} \le \quad \text{const.}[C(\delta) + \delta u(2L)]L^{2-d}D_\Lambda(\sqrt{f}) \tag{7.2}$$

under the assumption (7.1)

*Step 1.* We can decompose

$$\langle f; \tilde{g}_U(\bar{\eta}_U)\rangle_{\mu_{L,n}} = \langle f; \{\tilde{g}(\bar{\eta}_U) - w\bar{\eta}_U\}\rangle_{\mu_{L,n}} + w\langle f; \bar{\eta}_U\rangle_{\mu_{L,n}} := \Omega_1 + \Omega_2, \tag{7.3}$$

where $w = (\tilde{g}')(\rho^c_U)$ and $\rho^c_U = E^{\mu_{L,n}}[\bar{\eta}_U]$ is the density. The second term $\Omega_2$ in (7.3) can be bounded with the following Lemma 7.1.

**Lemma 7.1** *Suppose the mixing conditions (A1) and (A3) are satisfied. Then for any probability density $f$ and any two nearest neighbor sites $x, y$ with $|x - \Lambda^c| \ge L^{1-\epsilon}$ and $|x - y| \le 1$,*

$$\{E^f[\eta_x - \eta_y]\}^2 \le const.u(2L) \sum_{|b-x|\ge \ell} |b-x|^{-d-\epsilon} D_b(\sqrt{f}) + C(\ell) \sum_{|b-x|\le \ell} |b-x|^{-d-\epsilon} D_b(\sqrt{f}). \tag{7.4}$$

*Here $\ell$ is any integer, $C(\ell)$ is a constant depending on $\ell$ and $u(L)$ is defined in (3.1).*

**Proof**. Let $h = \eta_x - \eta_y$. Recall the definition (6.1) and the $\sigma$-algebra $\mathcal{F}_j$. Rewrite the left side of (7.4) as

$$E^{\mu_{L,n}}[h; f] = \sum_{j=0}^{\infty} E^{\mu_{L,n}}[(h_j - h_{j+1})f]$$

For each j fixed, from the definition of conditional expectation,

$$\langle f; h^{(j)} - h^{(j+1)}\rangle_{\mu_{L,n}} = E^{\mu_{L,n}}[E^{\mu_{L,n}}[f(h^{(j)} - h^{(j+1)})|\mathcal{F}_{j+1}]].$$

Consider the eigenvalue problem (with $k = \ell^{j+1}$)

$$\mathcal{E}(f) = \beta^{-1}k^{-1-d-\epsilon}[u(k)k^2 D_{j+1}(\sqrt{f})] - \int f(h^{(j)} - h^{(j+1)})d\mu_{j+1}$$

Here $\mu_{j+1} = \mu_{L,n}|_{\mathcal{F}_{j+1}}$ and $D_{j+1}$ is the Dirichlet form with respect to $\mu_{j+1}$. By definition of $u(k)$, we can replace $[u(k)k^2 D_{j+1}(\sqrt{f})]$ by the entropy $k^d s(f/\mu_{j+1})$ to have a lower bound. From the entropy bound (5.9) one can bound $\mathcal{E}(f)$ by

$$\mathcal{E}(f) \leq \beta^{-1}k^{-1-d-\epsilon}\log\int \exp[\beta k^{1+d+\epsilon}(h^{(j)} - h^{(j+1)})]d\mu_{j+1}$$

From (A3), one has

$$\| E^{\mu_{j+1}}[\eta_x - \eta_y] \|_\infty \leq \text{const.}k^{-(d/2)-1-\epsilon} \tag{7.5}$$

Hence

$$k^{1+\epsilon} \| h^{(j)} - h^{(j+1)} \|_\infty \leq k^{1+\epsilon} \| E^{\mu_{j+1}}[\eta_x - \eta_y] \|_\infty \leq \text{const.}k^{-(d/2)-\epsilon}$$

Since $\int f(h^{(j)} - h^{(j+1)})d\mu_{j+1} = 0$, together with Lemma 5.2 one can bound $\mathcal{E}(f)$ by $\beta k^{-1-\epsilon}$. To summarize, we have proved

$$\int f(h^{(j)} - h^{(j+1)})d\mu_{j+1} \leq \beta^{-1}k^{-1-d-\epsilon}[u(k)k^2 D_{j+1}(\sqrt{f})] + \beta k^{-1-\epsilon}$$

for all $\beta > 0$. Optimizing over $\beta$, we have

$$\{\int f(h^{(j)} - h^{(j+1)})d\mu_{j+1}\}^2 \leq k^{-d-\epsilon}[u(k)D_{j+1}(\sqrt{f})].$$

Summing over $j$ and divide the summation into $j = 1$ and $j > 1$, we have proved Lemma 7.1.

□

We now return to the proof of Lemma 6.2. For any constant $\sigma \ll \gamma$ rewrite $\Omega_2$ as $\Omega_2 = \Omega_3 + \Omega_4$ where

$$\begin{aligned}\Omega_3 &= C\langle f; Av_{x\in U,|x-\Lambda^c|\geq\sigma L}\eta_x\rangle_{\mu_{L,n}} = C\langle f; Av_{x\in U,|x-\Lambda^c|\geq\sigma L}Av_{y\in\Lambda}(\eta_x - \eta_y)\rangle_{\mu_{L,n}} \\ \Omega_4 &= C\sigma\langle f; Av_{x\in U,|x-\Lambda^c|\leq\sigma L}\ \eta_y\rangle_{\mu_{L,n}}\end{aligned} \tag{7.6}$$

We can decompose $\Omega_3 = \Omega_5 + \Omega_6$ with

$$\begin{aligned}\Omega_5 &= C\langle f; Av_{x\in U, |x-\Lambda^c|\geq \sigma L} Av_{y\in\Lambda, |y-\Lambda^c|\geq\sigma L}\ (\eta_x - \eta_y)\rangle_{\mu_{L,n}}, \\ \Omega_6 &= C\sigma\langle f; Av_{y\in\Lambda, |y-\Lambda^c|\leq\sigma L}\ \eta_y\rangle_{\mu_{L,n}}.\end{aligned} \tag{7.7}$$

From Schwarz's inequality, $\Omega_2$ (7.3) can be bounded by $\Omega_2^2 \leq \Omega_4^2 + \Omega_5^2 + \Omega_6^2$

Since $\Omega_4$ and $\Omega_6$ are of similar, we can absorb $\Omega_4$ into $\Omega_6$. Note that there is a factor $\sigma$ appearing in the definition of $\Omega_6$.

We shall bound $\Omega_4$ in the last step. We now bound $\Omega_5$. Connect the site $x$ to $y$ by a path $\gamma_{xy}$ with $|\gamma| \leq \text{const.}L$. There are many choices of $\gamma_{xy}$. For example, in dimension 2 one can fix a canonical choice by first connecting $x = (x_1, x_2)$ to $(x_1, y_2)$ via a straight line parallel to $y-$axis. Then connecting $(x_1, y_2)$ to $y = (y_1, y_2)$ via a straight line parallel to $x-$axis. Rewrite

$$\eta_x - \eta_y = \sum_{i=1}^{|\gamma|} \eta_i - \eta_{i+1}$$

where $|\gamma|$ is the length of $\gamma$ and we assume that $\gamma_1 = x, \gamma_{|\gamma|+1} = y$. By the Schwarz's inequality

$$\langle f; \eta_x - \eta_y\rangle^2_{\mu_{L,n}} \leq \text{const.}L \sum_{i=1}^{|\gamma|} \langle f; \eta_i - \eta_{i+1}\rangle^2_{\mu_{L,n}},$$

where we have bounded $|\gamma|$ by const.$L$. We can now apply lemma 7.1 to bound the last term. Summing over $x, y$, we can bound $\Omega_3$ by

$$\Omega_5^2 \leq [\text{const.} + \delta u(2L)] L^{2-d} D_\Lambda(\sqrt{f}) \tag{7.8}$$

*Step 3.* We now bound $\Omega_1$ (7.3). From the entropy bound (5.9), it is bounded by

$$\Omega_1 \leq \beta^{-1} L^{-d} \log E^{\mu_{L,n}} \left[\beta L^d (\tilde{g}_U(\bar{\eta}_U) - w(\bar{\eta}_U - \rho_U^c))\right] + \beta^{-1} u(2L) L^{2-d} D_\Lambda(\sqrt{f}) \tag{7.9}$$

where $\rho_U^c = E^{\mu_{L,n}}[\bar{\eta}_U]$ is a constant. From Theorem 5.6 the expectation in (7.9) can be bounded by

$$\beta^{-1} L^{-d} \log E^{\mu_{L,n}} \left\{ \exp\left[\beta L^d \{\tilde{g}_U(\bar{\eta}_U) - w(\bar{\eta}_U - \rho_U^c)\}\right] \right\} \leq \delta\beta$$

provided that $\beta \leq \beta_0 \ll 1$ (5.13). Here $\delta$ is a small constant. Optimizing over $\beta$ one has

$$\Omega_1^2 = \langle f; \tilde{g}_U(\bar{\eta}_U) - w(\bar{\eta}_U - \rho_U^c)\rangle_{\mu_{L,n}}^2 \leq \delta u(2L) L^{2-d} D_\Lambda(\sqrt{f}) \tag{7.10}$$

provided that

$$\beta = \delta^{-1/2} \left[ u(2L) L^{2-d} D_\Lambda(\sqrt{f}) \right]^{1/2} \leq \beta_0$$

The last bound holds by choosing $\delta_1$ in the assumption (7.1) sufficiently small. Hence combining (7.3), (7.8), (7.10) we have thus proved

$$\langle f; \tilde{g}_U(\bar{\eta}_U)\rangle_{\mu_{L,n}}^2 \leq [\text{const.} + \delta u(2L)] L^{2-d} D_\Lambda(\sqrt{f}) + \Omega_6^2 \tag{7.11}$$

with $\Omega_4$ defined in (7.7), provided that (7.1) holds. Note that the contribution for $x \in U, |x - \Lambda^c| \leq \sigma L$ is absorbed into $\Omega_4$. Combining with Lemma 6.1, we have proved

$$\begin{aligned} \langle f; g\rangle_{\mu_{L,n}}^2 \leq & \text{const.} \ell^{2d} u(2L) \sum_{|b-z| \leq L/\ell^{2d+2}} \Psi(|z-b|) D_b(\sqrt{f}) \\ & + 2[\text{const.} + \delta u(2L)] L^{2-d} D_\Lambda(\sqrt{f}) + 2\Omega_6^2 \end{aligned} \tag{7.12}$$

for any local function $g$.

*Step 4.* Finally we bound $\Omega_6$. Replacing $g$ in (7.12) by $\eta_z$ and averaging over $z \in \Lambda, |z - \Lambda^c| \leq \sigma L$, one has

$$\begin{aligned} \sigma^{-2}\Omega_6^2 \leq & \text{const.} \ell^{2d} u(2L)\ Av_{z \in \Lambda, |z-\Lambda^c| \leq \sigma L} \sum_{|b-z| \leq L/\ell^{2d+2}} \Psi(|z-b|) D_b(\sqrt{f}) \\ & + [\text{const.} + \delta u(2L)] L^{2-d} D_\Lambda(\sqrt{f}) + 2\Omega_6^2 \ ; \end{aligned}$$

here we have dropped unimportant numerical factors. The first term on the right side is bounded by $\text{const.} u(2L) L^{2-d} D_\Lambda(\sqrt{f})$. We have thus proved

$$\Omega_6^2 \leq [\text{const.} + \delta u(2L)] L^{2-d} D_\Lambda(\sqrt{f}).$$

Using this bound in (7.11) we have proved Lemma 6.2 assuming (7.1). Together with Lemma 6.3 we have proved Lemma 6.2. This concludes our proof of Theorem 2.1 except the proof of (A2-4) and Lemmas 5.4 and 5.5.

## VIII. Proof of Theorem 2.2 and Lemmas 5.4, 5.5

In this section we prove Lemma 5.4, 5.5 and Theorem 2.2. For the rest of this section we assume assumption (A1) holds for all $\lambda$ unless otherwise stated. We start with (A2) and Lemma 5.5 concerning only the grand canonical ensembles.

**Proof of (A.2)** First of all, notice that as $\rho \to 0$ one has $\lambda \to -\infty$. Hence the Gibbs state can be understood as a perturbation of the independent measure by the Hamiltonian $H$. If $H = 0$, one has a independent measure and (A2) can be checked directly. If $\rho \leq \epsilon$ or $\rho \geq 1-\epsilon$, (A2) follows from the standard cluster expansion [R] and we omit the details. We now consider the case $\rho$ is bounded away from 0 or 1.

The derivatives of $\tilde{g}$ can be computed as:

$$\partial_y \hat{g}(y) = \partial_\lambda \hat{g}(y) \frac{\partial \lambda}{\partial y} = \langle g; \sum_x \eta_x \rangle_{\mu_{L,\lambda}} / \langle \sum_x \eta_x ; A v_x \eta_x \rangle_{\mu_{L,\lambda}}$$

$$\partial_y^2 \hat{g}_U(y) = \partial_\lambda^2 \hat{g}_U(y) \Big(\frac{\partial \lambda}{\partial y}\Big)^2 - \partial_\lambda \hat{g}_U(y) \Big(\frac{\partial^2 y}{\partial \lambda^2}\Big)\Big(\frac{\partial \lambda}{\partial y}\Big)^3$$
$$= \langle \sum_x \eta_x ; A v_x \eta_x \rangle_{\mu_{L,\lambda}}^{-3} \times$$
$$\Big[\langle g; \sum_x \eta_x ; \sum_x \eta_x \rangle_{\mu_{L,\lambda}} \langle \sum_x \eta_x ; A v_x \eta_x \rangle_{\mu_{L,\lambda}} - \langle g; \sum_x \eta_x \rangle_{\mu_{L,\lambda}} \langle A v_x \eta_x ; \sum_x \eta_x ; \sum_x \eta_x \rangle_{\mu_{L,\lambda}}\Big]$$

Here the expectation is with respect to the grand canonical Gibbs state $\mu_{U,\lambda}$ with $\lambda$ chosen to give the correct density $\rho$. From (A.1) for $\epsilon \leq \rho \leq 1-\epsilon$ the covariances can be bounded by

$$\langle A v_x \eta_x ; \sum_x \eta_x ; \sum_x \eta_x \rangle_{\mu_{L,\lambda}} \leq \text{const.}$$
$$\langle g; \sum_x \eta_x \rangle_{\mu_{L,\lambda}} \leq \text{const.}, \quad \langle g; \sum_x \eta_x ; \sum_x \eta_x \rangle_{\mu_{L,\lambda}} \leq \text{const.}$$

Furthermore, we claim the following bound on the compressibility holds

$$\langle \sum_x \eta_x ; A v_x \eta_x \rangle_{\mu_{L,\lambda}} \geq C_\rho$$

for some constant $C_\rho$ depending on $\rho$. Hence $\hat{g}_U$ satisfies (A2) in this region assuming this bound on the compressibility.

Finally we have to prove the bound on the compressibility. Recall the definition of $\sigma_x$ from (2.8). For any local functions $f$ and $h$,

$$\Big\langle [f(\sigma_x\eta)-f(\eta)]\ [h(\sigma_x\eta)-h(\eta)]\Big\rangle = -\Big\langle f(\eta)\ [\exp\{-H(\sigma_x\eta)+H(\eta)\}+1]\ [h(\sigma_x\eta)-h(\eta)]\Big\rangle$$

Let

$$g = -\,[\exp\{-H(\sigma_0\eta)+H(\eta)\}+1]\ [h(\sigma_0\eta)-h(\eta)], \quad h(\eta)=\eta_0.$$

Denote by $g_x$ the translation of $g$ to $x$. Recall the range of interactions in the Hamiltonian is one. Let $\Lambda^0$ denote the interior of $\Lambda$ defined by $\Lambda^0=\{x\in\Lambda : |x-\Lambda^c|>2\}$. From the Schwarz inequality,

$$\Big\langle \Big\{L^{d/2}Av_{x\in\Lambda}(\eta_x-\rho)\Big\}^2\Big\rangle_{\mu_{L,\lambda}} \ge \Big[\Big\langle Av_{x\in\Lambda^0}\, g_x, \sum_x(\eta_x-\rho)\Big\rangle_{\mu_{L,\lambda}}\Big]^2 \Big[L^d\Big\langle\ \{Av_{x\in\Lambda^0}\, g_x\}^2\Big\rangle_{\mu_{L,\lambda}}\Big]^{-1}$$

By the definition of $g$,

$$\Big\langle Av_{x\in\Lambda^0}g_x, \sum_x(\eta_x-\rho)\Big\rangle_{\mu_{L,\lambda}} = Av_{x\in\Lambda^0}\Big\langle(1-2\eta_x)^2\Big\rangle_{\mu_{L,\lambda}} = 1$$

Here we have used $(1-2\eta_x)^2=1$ since $\eta_x\in\{0,1\}$. From the mixing condition, there is a constant $C_\rho$ such that

$$L^d\Big\langle\ \{Av_{x\in\Lambda^0}\, g_x\}^2\Big\rangle_{\mu_{L,\lambda}} \le C_\rho$$

We have thus proved that

$$\Big\langle \Big\{L^{d/2}Av_{x\in\Lambda}(\eta_x-\rho)\Big\}^2\Big\rangle_{\mu_{L,\lambda}} \ge C_\rho^{-1}$$

This proves the lower bound on the compressibility and concludes (A2). □

**Proof of Lemma 5.5.** Recall the identity

$$\langle X;X|\Lambda\setminus U\rangle = \langle X^2|\Lambda\setminus U\rangle - \langle X|\Lambda\setminus U\rangle^2$$

and $\langle X|\Lambda\setminus U\rangle$ denotes the conditional expectation of $X$ with respect to the $\sigma$-algebra generated by the configuration in $\Lambda\setminus U$. Hence the correlation functions appearing in Lemma 5.5 can be decomposed as

$$\langle\bar\eta_U;\bar\eta_U\rangle_{\mu_{L,\theta,\lambda}} = \langle\bar\eta_U;\bar\eta_U\mid\Lambda\setminus U\rangle_{\mu_{L,\theta,\lambda}} + \langle E^{\mu_{L,\theta,\lambda}}[\bar\eta_U\mid\Lambda\setminus U]; E^{\mu_{L,\theta,\lambda}}[\bar\eta_U\mid\Lambda\setminus U]\rangle_{\mu_{L,\theta,\lambda}}$$

From the mixing assumption (A1), the first term on the right side is bounded by $L^{-d}$. For the second term, again from the mixing condition, the spectral gap of the Glauber dynamics [MO,LY,SZ] is bounded by a universal constant. Hence one has

$$\langle E^{\mu_{L,\theta,\lambda}}[\bar\eta_U \mid \Lambda\backslash U]; E^{\mu_{L,\theta,\lambda}}[\bar\eta_U \mid \Lambda\backslash U]\rangle_{\mu_{L,\theta,\lambda}} \le \sum_{x\in\Lambda\backslash U, |x-U|\le 1} E^{\mu_{L,\theta,\lambda}}[(\sigma_x E^{\mu_{L,\theta,\lambda}}[\bar\eta_U \mid \Lambda\backslash U])^2]$$

where $\sigma_x$ is defined in (2.8). From the mixing assumption,

$$\sigma_x E^{\mu_{L,\theta,\lambda}}[\bar\eta_U \mid \Lambda \setminus U] \le \text{const}.L^{-d}$$

Hence

$$\sum_{x\in\Lambda\backslash U, |x-U|\le 1} E^{\mu_{L,\theta,\lambda}}[(\sigma_x E^{\mu_{L,\theta,\lambda}}[\bar\eta_U \mid \Lambda \setminus U])^2] \le \text{const}.L^{-d}\,.$$

This concludes Lemma 5.5. $\square$

We now prove (A3-4) and Lemma 5.4. Recall that $\mu_{L,\lambda}$ is the Gibbs state on a cube $\Lambda_L$ with chemical potential $\lambda$ and some fixed boundary condition. Let $N = L^d$. Recall $U$ is a subcube of size $L/\gamma$ with the property that $\Lambda_L \setminus U$ is "fat" (5.1). Let $\mu_{L,\theta,\lambda}$ be the probability measure with density relative to $\mu_{L,\lambda=0}$ given by

$$\exp(\theta \sum_{x\in U} \eta_x + \lambda \sum_{x\in\Lambda_L} \eta_x\,)\, Z(\theta,\lambda)^{-1} \tag{8.1}$$

where $Z(\theta,\lambda)$ is the normalization. Clearly, when $\theta = 0$ the measure $\mu_{L,\theta=0,\lambda}$ reduces to $\mu_{L,\lambda}$. We have proved that $\mu_{L,\theta,\lambda}$ satisfies the mixing condition (A1) for all $\theta$ and $\lambda$ and for all cubes if $\mu$ satisfies the mixing condition (A1) for all $\lambda$ and for all cubes.

We now state our main result Theorem 8.1, a local limit theorem to be proved in the section IX.

**Theorem 8.1 (Local Limit Theorem)** *Suppose the mixing condition (A1) holds. Let $L^{-d}E^{\mu_{L,\theta,\lambda}}\sum_x \eta_x = \rho_\theta$ is the density. Let $X = L^{-d/2}\sum_x(\eta_x - \rho_\theta)$, and let $\sigma_\theta^2$ be the variance defined by $\sigma_\theta^2 = E^{\mu_{L,\theta,\lambda}}X^2$. Let $\psi_{L,t}$ be the density of $X_\theta$. Suppose the density satisfies*

$$L^{-d/2-\epsilon} \le \rho \le 1 - L^{-d/2-\epsilon}.$$

*Then there is a universal constant of order one such that $g_{L,\theta}$ satisfies*

$$\psi_{L,\theta}(x) = (2\pi)^{-1/2} C_L \sigma_\theta^{-1} \exp[-x^2/(2\sigma_\theta^2)] \left[1 + N^{-1/2} p_1(x) + O(L^{-d+\epsilon}\rho^{-1})\right] \tag{8.2}$$

*for some $\epsilon > 0$. Here*

$$p_1(x) = \mu_3 H_3/6,\ \mu_3 = \sigma_\theta^{-3} E^{\mu_{L,\theta,\lambda}} X_\theta^3$$

*and $H_3$ is the Hermite polynomial of degree 3:*

$$H_3(x) = x^3 - 3x$$

We shall now represent $\psi_{L,\theta}$ using the Fourier inversion formula. Recall the following elementary identity for Fourier series,

$$L^{-d} \sum_{k=0}^{N-1} e^{i2\pi kx/N} = \begin{cases} 1 & \text{if } x = 0\ ; \\ 0 & \text{otherwise.} \end{cases} \tag{8.3}$$

where

$$x \in \{-N+1, \ldots, 0, 1, \ldots, N-1\}.$$

Hence the canonical measure $\mu_{L,n}$ can be represented as

$$d\mu_{L,n} = L^{-d} \sum_{k=0}^{N-1} e^{i2\pi k(\bar\eta - (n/N))} d\mu_{L,\lambda}/Z_{L,n},$$

where

$$Z_{L,n} = E^{\mu_{L,\lambda}} \left[L^{-d} \sum_{k=0}^{N-1} e^{i2\pi k(\bar\eta - (n/N))}\right]$$

and the chemical potential $\lambda$ is chosen such that $E^{\mu_{L,\lambda}}[\bar\eta] = \rho$. Hence

$$\psi_{L,\theta}(y) = L^{-d} \sum_{k=0}^{N-1} g_{N,\theta}^{(k)}(y) = L^{-d} \sum_{k=0}^{N-1} E^{\mu_{L,\theta,\beta}} \left[\exp\{i2\pi k L^{-d} [\sum_x \eta_x - y]\}\right]. \tag{8.4}$$

with $\beta$ chosen such that

$$E^{\mu_{L,\theta,\beta}}[\bar\eta] = \rho. \tag{8.5}$$

**Proof of Lemma 5.4** By definition of $\phi$ in Lemma 5.4 ,

$$\exp[(L^d/\gamma^d)\phi(\theta,\rho)] = E^{\mu_{L,n}}[\exp[\theta\sum_{x\in U}\eta_x]]$$

$$= \frac{E^{\mu_{L,\lambda=0}}\left[\exp\{\beta\sum_{x\in\Lambda}\eta_x + \theta\sum_{x\in U}\eta_x\}\mathbf{1}(\bar\eta=\rho)\right]}{E^{\mu_{L,\lambda=0}}\left[\exp\{\beta\sum_{x\in\Lambda}\eta_x + \theta\sum_{x\in U}\eta_x\}\right]}\exp[(L^d/\gamma^d)p(\theta,\rho)]$$

$$\times\left\{\frac{E^{\mu_{L,\lambda=0}}\left[\exp\{\lambda\sum_{x\in\Lambda}\eta_x\}\mathbf{1}(\bar\eta=\rho)\right]}{E^{\mu_{L,\lambda=0}}\left[\exp\{\lambda\sum_{x\in\Lambda}\eta_x\}\right]}\exp[(L^d/\gamma^d)p(\theta=0,\rho)]\right\}^{-1}$$

with $\beta$ satisfying (8.5) and $E^{\mu_{L,\lambda}}[\bar\eta]=\rho$. Hence

$$\phi(\theta,\rho) = p(\theta,\rho) - p(0,\rho) + (\gamma^d/L^d)[\log\psi_{L,\theta}(0) - \log g_{L,\theta=0}(0)] \tag{8.6}$$

From Theorem 8.1, $\log\psi_{L,\theta}(0)$ is bounded by some constant independent of $L$ if $\rho\ge L^{-d/2}$. This concludes Lemma 5.4. $\Box$.

**Proof of (A4)** We first state the following corollary of Theorem 8.1.

**Corollary 8.2.** *Suppose $\mu_{L,n}$ is the measure defined in (8.11). Suppose $u$ is a local function. Then*

$$E^{\mu_{L,n}}[u] = E^{\mu_{L,\lambda}}[u] + O(L^{-d+\epsilon}\rho^{-1}). \tag{8.7}$$

*Here the expectation is with respect to the measure $\mu_{L,\lambda}$ with $\lambda$ chosen to give the correct density, i.e., $E^{\mu_{L,\lambda}}[\bar\eta]=\rho$.*

We shall prove this corollary at the end of this section. Assuming this corollary, we immediately have (A4) if $\rho\ge L^{-d/2+2\epsilon}$. For any local function $u$ we have

$$|E^{\mu_{L,n}}[u(\eta)] - u(\eta=0)| \le \text{const.}\rho$$

and similarly if $\mu_{L,n}$ is replaced by $\mu_{L,\lambda}$. Hence (A4) holds trivially for $\rho\le L^{-d/2-\epsilon}$. We now prove (A4) assuming $L^{-d/2-\epsilon}\le\rho\le L^{-d/2+\epsilon}$.

From the local limit theorem 8.1 with $\theta=0$, we obtain a large deviation estimate via standard Cramer method. Note that the variance $\sigma$ satisfies $\sigma^2\le\text{const.}\rho$. Hence we have the following large deviation estimate

$$E^{\mu_{L,\lambda}}\left[|\bar{\eta}-\rho| \geq L^{-d/2+\epsilon}\rho^{1/2}\right] \leq \exp[-CL^{\epsilon}] \tag{8.8}$$

Since this estimate is standard and can be proved using arguments similar to the proof of Lemma 5.3, we omit the detail.

Recall (4.10)

$$\begin{aligned}
&E^{\mu_{L,n}}[u] - E^{\mu_{L,n+1}}[u] \\
&= E^{\mu_{L,n}}[\bar{F}]^{-1}E^{\mu_{L,n}}[u;\bar{F}] + (L^d/n+1)Av_{x\in\Lambda}E^{\mu_{L,n+1}}[(u(\sigma_x\eta)-u(\eta))\eta_x].
\end{aligned}$$

We have changed notations to our setting and $F$ is defined in (4.4). The last term is bounded by $CL^{-d}$ with the constant independent of $\rho$. Also, from definition of $F$, we have $E^{\mu_{L,n}}[\bar{F}]^{-1} \leq C$ (cf: proof of Lemma 4.2). Suppose $u$ is a local function at $z$ with range $S$. The variance can be decomposed as

$$E^{\mu_{L,n}}[u;\bar{F}] = L^{-d}\sum_{|x-z|\leq S+2}E^{\mu_{L,n}}[u;F_x] + L^{-d}\sum_{|x-z|>S+2}E^{\mu_{L,n}}[u;F_x]$$

Since the range of the interaction in the Hamiltonian is one and the density is small, one can check that

$$|E^{\mu_{L,n}}[u;F_x]| \leq \begin{cases} C\rho & \text{if } |x-z|\leq S+2\ , \\ C\rho^2 & \text{if } |x-z|>S+2 \end{cases}$$

Hence

$$|E^{\mu_{L,n}}[u;\bar{F}]| \leq C\rho/L^d + C\rho^2 \leq CL^{-d+2\epsilon}$$

We have thus proved that

$$|E^{\mu_{L,n}}[u] - E^{\mu_{L,n+1}}[u]| \leq CL^{-d+2\epsilon}$$

Together with (8.8), we have

$$\begin{aligned}
&|E^{\mu_{L,\lambda}}[u] - E^{\mu_{L,n}}[u]| \\
&\leq \sup_{|m-n|\leq CL^{d/2}\rho^{1/2}} |E^{\mu_{L,n}}[u] - E^{\mu_{L,m}}[u]| \\
&\leq CL^{d/2}\rho^{1/2}L^{-d+2\epsilon} \leq CL^{-d/2-\epsilon}
\end{aligned}$$

This proves (A4). $\square$.

**Proof of (A.3)** From Corollary 8.2, we have proved (A.3) if the dimension $d \geq 3$ and $L^{-\epsilon} \leq \rho \leq 1 - L^{-\epsilon}$. The case $\rho \leq L^{-\epsilon}$ can be checked by standard low density expansion [R]. For $d \leq 2$, one can carry out the local limit theorem, Theorem 8.1, to the next order and hence prove (A.3) . The proof will be somehow complicated and will not be presented here. An alternative approach will be given in a forthcoming paper [VY].

**Proof of Corollary 8.2** Without loss of generality We can assume that $E^{\mu_{L,n}}[u] = 0 = E^{\mu_{L,n}}[X] = y$. Our goal is to compute $E^{\mu_{L,\lambda}}[u\mathbf{1}(X=0)]/E^{\mu_{L,\lambda}}[\mathbf{1}(X=0)]$. Since $u$ may not be positive, we can not apply Theorem 8.1 to $E^{\mu_{L,\lambda}}[\cdot u]$. Let us assume that $1+u>0$ for simplicity. Hence $(u+1)\mu_{L,\lambda}$ is a probability measure. Define

$$X_u = N^{-1/2}\sum_x \eta_x - z, \quad z = E^{\mu_{L,\lambda}}[Xu] = \langle X; u\rangle - \mu_{L,\lambda}. \tag{8.9}$$

Apply Theorem 8.1 to the new probability measure $(u+1)\mu_{L,\lambda}$,

$$E^{\mu_{L,\lambda}}[(1+u)\mathbf{1}(X=0)] = E^{(1+u)\mu_{L,\lambda}}[\mathbf{1}(X_u=-z)] = \psi_L^{(1+u)}(z)$$

where $\psi_L^{(1+u)}$ is the function obtained by replaced all expectation in Theorem 8.1 from with respect to $\mu_{L,\theta,\lambda}$ to with respect to $(1+u)\mu_{L,\lambda}$. By definition,

$$E^{\mu_{L,\lambda}}[u\mathbf{1}(X=0)]/E^{\mu_{L,\lambda}}[\mathbf{1}(X=0)] = \psi_L^{(1+u)}(-z)\psi_L(0) - 1$$

The variance of $X$ with respect to the measure $(1+u)\mu_{L,\lambda}$ is given by

$$\sigma_{1+u}^2 = \langle (X-z)^2(1+u)\rangle_{\mu_{L,\lambda}} = \langle X^2(1+u)\rangle_{\mu_{L,\lambda}} - z^2 = \sigma^2 + \langle X^2 u\rangle_{\mu_{L,\lambda}} - z^2\,.$$

From Theorem 8.1, we have

$$\psi_L(0) = (2\pi)^{-1/2} C_L \sigma^{-1}\left[1 + O(L^{-d+\epsilon}\rho^{-1})\right]$$

From (8.9) and the mixing condition (A.1), $z = O(L^{-d/2}\rho)$. Applying Theorem 8.1 to the measure $(1+u)\mu_{L,\lambda}$ and using $z = O(L^{-d/2}\rho)$, we have

$$\psi_L^{(1+u)}(-z) = (2\pi)^{-1/2} C_L \sigma_{1+u}^{-1} \exp\left[-z^2/(2\sigma_{1+u}^2)\right]\left[1 + O(L^{-d+\epsilon}\rho^{-1})\right]$$

Combining these two estimates, one has

$$\psi_L^{(1+u)}(z)\psi_L(0)^{-1} - 1 = \frac{\sigma}{\sigma_{1+u}} \exp\{-\frac{z^2}{2\sigma_{1+u}^2}\} \left[1 + O(L^{-d+\epsilon}\rho^{-1})\right] - 1.$$

From the definition of $u$ and $z = O(L^{-d/2}\rho)$,

$$\frac{\sigma}{\sigma_{1+u}} \exp\left[-\frac{z^2}{2\sigma_{1+u}}\right]$$
$$= \left[1 + \frac{\langle X^2 u\rangle_{\mu_{L,\lambda}} + z^2}{\sigma^2}\right]^{-1/2} \left[1 - \frac{z^2}{2\left(\sigma^2 + \langle X^2 u\rangle_{\mu_{L,\lambda}} - z^2\right)}\right] = 1 + O(L^{-d+\epsilon}\rho^{-1})$$

We have thus proved

$$\psi_L^{(1+u)}(z)\psi_L(0)^{-1} - 1 \le O(L^{-d+\epsilon}\rho^{-1})$$

This concludes the proof of Corollary 8.2. □

## IX. Local Limit Theorem for Gibbs Measures

We now prove the local limit theorem 8.1. Our method is straightforward and based on a martingale decomposition which helps in the organization of error terms. Martingale methods have a long history; our proof is not novel .

Recall the identity (8.4)

$$\psi_{L,\theta}(y) = L^{-d} \sum_{k=0}^{L^d-1} \psi_{L^d,\theta}^{(k)}(y) = L^{-d} \sum_{k=0}^{L^d-1} E^{\mu_{L,\theta,\lambda}} \left[\exp\{i2\pi k L^{-d} \sum_x (\eta_x - y)\}\right]. \tag{9.1}$$

By periodicity, for $L^d$ odd,

$$L^{-d} \sum_{k=0}^{L^d-1} e^{i2\pi k(\bar\eta - y)} = \frac{1}{2} L^{-d} \sum_{|k|<L^d/2} e^{i2\pi k(\bar\eta-y)}. \tag{9.2}$$

Hence we shall replace the summation in (9.1) with (9.2).

**Step 1. Cutoff for $k > \rho^{-1/2}L^{d/2+\epsilon}$.** Let

$$\Omega^{(\epsilon)} = L^{-d} \sum_{L^{d/2+\epsilon}<|k|<L^d/2} e^{i2\pi k(\bar\eta-y)}$$

Recall the range of interaction in the Hamiltonian is one. Let $Q = \{x \in \Lambda_L : x/2 \in Z^d\}$ be a sublattice of $\Lambda$ and let $\Gamma = \Lambda_L \setminus Q$. Denote the expectation of $\eta_x$ conditioned on $\{\eta_y, y \in \Gamma\}$ by $p_x^\Gamma = E^{\mu_{L,\theta,\lambda}}[\eta_x | \eta_y, y \in \Gamma]$ . Let $p_x$ be the expectation of $\eta_x$ with respect to the independent measure

$$q_{\theta,\lambda} = \exp[\theta \sum_{x \in U} \eta_x + \lambda \sum_{x \in \Lambda_L} \eta_x] / \text{ normalization}$$

Since the interaction between $\eta_x$ and its neighbors is uniformly bounded for each $x$ fixed, there is a constant $C$ such that

$$C^{-1} \le p_x / p_x^\Gamma \le C$$

Note that this bound is independent of the shape of $\Gamma$. It is very easy to compute $p_x$ and thus bounds on $p_x^\Gamma$ independent of the configuration of $\Gamma$ can be obtained. By definition, $p_x$ takes only two values, say $a$ and $b$, depending on whether $x$ belongs to $U$ or not. Let $p_x = a$ if $x \in U$ and $p_x = b$ otherwise. Since the total density is $\rho$, one has

$$a|U| + b(L^d - |U|) = L^d \rho$$

This implies that $\rho \le a + b \le \text{const.}\rho$ since $L^d/|U|$ is bounded (depends on $\gamma$). Hence we have

$$C'\rho \le p_x + p_x^\Gamma \le C\rho. \tag{9.3}$$

Since $\eta_x$ are independent random variables after conditioning on $\Gamma$, one has

$$\begin{aligned} E^{\mu_{L,\theta,\lambda}}\big[\prod_{x \notin \Gamma} e^{i2\pi k \eta_x L^{-d}} \mid \eta_y, y \in \Gamma\big] &= \prod_{x \notin \Gamma} E^{\mu_{L,\theta,\lambda}}\big[e^{i2\pi k \eta_x L^{-d}} \mid \eta_y, y \in \Gamma\big] \\ &= \prod_{x \notin \Gamma} [p_x^\Gamma e^{i2\pi k L^{-d}} + (1 - p_x^\Gamma)], \end{aligned}$$

Recall the following elementary bound: For any $0 \le a \le 1$ and $-\pi \le \theta \le \pi$ one has

$$|ae^{i\theta} + 1 - a|^2 = 1 - 2a(1-a)(1 - cos\theta) \le \exp(-a(1-a)\theta^2)$$

We have thus proved

$$
\begin{aligned}
|E^{\mu_{L,\theta,\lambda}}[\Omega^{(\epsilon)} \mid \Gamma]| &\le L^{-d} \sum_{L^{d/2+\epsilon}<|k|<L^d/2} \exp\big[-\sum_{x\in Q} p_x^\Gamma k^2 L^{-2d}\big] \\
&\le L^{-d} \sum_{L^{d/2+\epsilon}<|k|<L^d/2} \exp(-\text{const.}\rho k^2 L^{-d})
\end{aligned}
$$

where we have used (9.3). From the range of $k$ one has

$$
|E^{\mu_{L,\theta,\lambda}}[\Omega^{(\epsilon)} \mid \Gamma]| \le \exp(-\text{const.}L^\epsilon).
$$

Therefore, the contribution of this region is negligible.

**Step 2. Perturbation Expansion** From the cutoff of $k$ and the Fourier inversion formula, our goal is to compute

$$
\psi_{L,\theta}^{(k)}(y) = E^{\mu_{L,\theta,\lambda}}\Big[\exp\{i2\pi k L^{-d}\sum_x(\eta_x-\rho)\}\Big]
$$

for $k$ satisfying

$$
|k| \le \rho^{-1/2} L^{d/2+\epsilon}.
$$

The following lemma is the key input.

**Lemma 9.1** *With the same assumption as in the local limit theorem, one has, for* $|k| \le \rho^{-1/2} L^{d/2+\epsilon}$,

$$
\begin{aligned}
\psi_{L,\theta}^{(k)}(y) &= E^{\mu_{L,\theta,\lambda}}\Big[\exp\{i2\pi k L^{-d}\sum_x(\eta_x-y)\}\Big] \\
&= \exp\big[\frac{-L^d t^2}{2}\langle X^2\rangle - itL^d(y-\rho)\big]\big\{1-\frac{it^3}{3}\mu_3 + O(L^{-d+\epsilon}\rho^{-1})\big\}
\end{aligned}
\tag{9.4}
$$

*where*

$$
t = 2\pi k L^{-d}, \quad X = L^{-d/2}\sum_x(\eta_x-\rho). \tag{9.5}
$$

Returning to the proof of Theorem 8.1. From the bound in step 1 we can estimate $\psi_{L,\theta}(y)$ by

$$\begin{aligned}\psi_{L,\theta}(y) &= \frac{1}{2}L^{-d} \sum_{|k|<L^d/2} \psi^{(k)}_{L,\theta}(y) \\ &= \frac{1}{2}L^{-d} \sum_{|k|<\rho^{-1/2}L^{d/2+\epsilon}} E^{\mu_{L,\theta,\lambda}}\Big[\exp\{i2\pi kL^{-d}\sum_x(\eta_x-\rho)\}\Big] + O(\exp(-\text{const}.L^\epsilon)).\end{aligned}$$

From Lemma 9.1,

$$\begin{aligned}\psi_{L,\theta}(y) = \frac{1}{2}L^{-d} \sum_{|k|<\rho^{-1/2}L^{d/2+\epsilon}} \exp\big[\frac{-L^d t^2}{2}\langle X^2\rangle - itL^d(y-\rho)\big]\times \\ \big\{1-\frac{it^3}{3}\mu_3 + O(L^{-d+\epsilon}\rho^{-1})\big\} + O(\exp(-\text{const}.L^\epsilon)).\end{aligned}$$

The summation in (9.4) can be approximated by integration. Instead of estimating the difference between the summation and integration, we use an universal constant $C_L$ to characterize their difference. Performing the "Gaussian summation", we conclude Theorem 8.1. □

**Proof of Lemma 9.1 Step 1** We shall prove Lemma 9.1 by induction and the martingale decomposition. First let us introduce the martingale. For simplicity of notation we assume $d=2$. Let $\Gamma$ be the cube of size $L^\epsilon$ in the upper right corner of $\Lambda$. We shall denote all configuration in this cube by $\eta_0$. We now define an order starting from the site in $\Lambda\setminus\Gamma$ right next to the lower left corner of $\Gamma$ as $x_1$. We then continue the order by wrapping around $\Gamma$. When this is done, we start the same procedure again but with $\Gamma$ enlarged by the sites already ordered. Continue this procedure, we have an order. For simplicity we denote $\eta_{x_j}$ by $\eta_j$. Let $\mathcal{F}_j$ be the $\alpha$-algebra generated by $\eta_i, i>j$. Let $W=\sum_{x\in\Lambda}\eta_x$. Define

$$W_j = E^{\mu_{L,\theta,\lambda}}[W \mid \mathcal{F}_j] - E^{\mu_{L,\theta,\lambda}}[W \mid \mathcal{F}_{j+1}]$$

Clearly, $W=\sum_{j=0}W_j$. From the mixing condition (A.1)

$$|W_j| \le \text{const}. \tag{9.6}$$

To prove this, rewrite $W_j$ as

$$W_j = \Big\{ \sum_{x \in \alpha_j} \eta_x - E^{\mu_{L,\theta,\lambda}}[\sum_{x\in\alpha_j} \eta_x | \mathcal{F}_{j+1}] \Big\}$$
$$+ \Big\{ E^{\mu_{L,\theta,\lambda}}[\sum_{x \in \alpha_i, i<j} \eta_x | \mathcal{F}_j] - E^{\mu_{L,\theta,\lambda}}[\sum_{x \in \alpha_i, i<j} \eta_x | \mathcal{F}_{j+1}] \Big\}$$

Clearly, the terms inside the first parenthesis satisfies the bound in (9.6) due to the size of subcubes chosen previously. To bound the second term, denote $E^{\mu_{L,\theta,\lambda}}[\sum_{x\in\alpha_i, i<j} \eta_x|\mathcal{F}_j]$ by $Z$. From the mixing assumption (A.1), the second term satisfies the same bound. This proves (9.6).

Rewrite $\psi^{(k)}_{L,\theta}(y)$ as

$$\begin{aligned} \psi^{(k)}_{L,\theta}(y) &= E^{\mu_{L,\theta,\lambda}}\Big[ \exp\{i2\pi k L^{-d} \sum_j W_j\} \Big] \\ &= E^{\mu_{L,\theta,\lambda}}\Big\{ \exp\{i2\pi k L^{-d} \sum_{i>0} W_i\} E^{\mu_{L,\theta,\lambda}}\big[\exp\{i2\pi k L^{-d} W_0 \mid \mathcal{F}_1\big] \Big\} \end{aligned} \tag{9.7}$$

Then the expectation with $\mathcal{F}_1$ given can be easily computed since, for $k$in the range we are interested in, the exponent

$$kL^{-d} W_0 \le \rho^{-1/2} L^{-d/2+2\epsilon} \ll 1$$

is small (recall $\rho \ge L^{-d/2+\epsilon}$). Thus one can simply expand the exponential to have

$$\begin{aligned} &E^{\mu_{L,\theta,\lambda}}\big[\exp\{i2\pi k L^{-d} W_0\} \mid \mathcal{F}_1\big] \\ &= 1 + it E^{\mu_{L,\theta,\lambda}}[W_0 \mid \mathcal{F}_1] - \frac{1}{2} E^{\mu_{L,\theta,\lambda}}[(tW_0)^2 \mid \mathcal{F}_1] \\ &\quad - \frac{i}{3!} E^{\mu_{L,\theta,\lambda}}[(tW_0)^3 \mid \mathcal{F}_1] + O(\rho^{-1} L^{-2d+8\epsilon}), \quad t = 2\pi k L^{-d} \end{aligned}$$

Note that the error term gain a factor of $\rho$ from taking the expectation. The first order term $E^{\mu_{L,\theta,\lambda}}[W_0 \mid \mathcal{F}_1] = 0$ by definition of $W_0$. The variance and the third moment are bounded by

$$E^{\mu_{L,\theta,\lambda}}[(W_0)^i \mid \mathcal{F}_1] \le \text{const.} L^\epsilon E^{\mu_{L,\theta,\lambda}}[\bar{\eta}_0 \mid \mathcal{F}_1], \qquad i = 2, 3.$$

Here $\bar{\eta}_0$ denote the density in the cube $\alpha_0$. From (9.3), we have

$$E^{\mu_{L,\theta,\lambda}}[\bar{\eta}_0 \mid \mathcal{F}_1] \leq \text{const.}\rho \ .$$

Using this bound, we have that the variance and the third moment are bounded by

$$E^{\mu_{L,\theta,\lambda}}[(W_0)^i \mid \mathcal{F}_1] \leq \text{const.}L^\epsilon \rho \qquad i = 2, 3.$$

Hence

$$\begin{aligned}
&E^{\mu_{L,\theta,\lambda}}\Big[\exp\{i2\pi k L^{-d} W_0\} \mid \mathcal{F}_1\Big] \\
&= 1 - \frac{t^2}{2}V_0 - \frac{i}{3!}E^{\mu_{L,\theta,\lambda}}\ \big[(tW_0)^3 \mid \mathcal{F}_1\big] + O(\rho^{-1}L^{-2d+8\epsilon}) \\
&= \exp\Big\{-\frac{t^2}{2}V_0\Big\}\left\{1 - \frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}[W_0^3 \mid \mathcal{F}_1] + O(\rho^{-1}L^{-2d+8\epsilon})\right\}.
\end{aligned}$$

where

$$V_0 = E^{\mu_{L,\theta,\lambda}}[W_0^2|\mathcal{F}_1] \leq \text{const.}L^\epsilon\rho \ . \tag{9.8}$$

Together with (9.7),

$$\begin{aligned}
\psi_{L,\theta}^{(k)}(y) = E^{\mu_{L,\theta,\lambda}}\Bigg\{ &\exp\{i2\pi k L^{-d}\sum_{i>0} W_i\}\exp[-\frac{t^2}{2}V_0] \\
&\times\left[1 - \frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}[W_0^3 \mid \mathcal{F}_1] + O(\rho^{-1}L^{-2d+8\epsilon})\right]\Bigg\}
\end{aligned} \tag{9.9}$$

We shall prove (9.4) inductively. Let us focus on the next term $W_1$ and condition on $\mathcal{F}_2$. Denote

$$B = itW_1 - \frac{t^2}{2}\Big\{V_0 - E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\Big\}$$

Hence $\exp[itW_1 - \frac{t^2}{2}V_0] = \exp[B - \frac{t^2}{2}E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]]$. From (9.6), (9.8) and the cutoff on $k$, $|t| = 2|\pi k/L^d| \leq \rho^{-1}L^{-d+2\epsilon}$. Hence

$$\begin{aligned}
&t|W_1| \leq tL^\epsilon \leq \rho^{-1/2}L^{-d/2+2\epsilon}, \\
&\frac{t^2}{2}\Big|V_0 - E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\Big| \leq \rho^{-1}L{-}d + 4\epsilon, \\
&|t^3|\Big|E^{\mu_{L,\theta,\lambda}}[W_0^3 \mid \mathcal{F}_1] - E^{\mu_{L,\theta,\lambda}}[W_0^3 \mid \mathcal{F}_2]\Big| \leq \rho^{-3/2}L{-}3d/2 + 6\epsilon
\end{aligned}$$

Expanding the exponential involving $B$,

$$E^{\mu_{L,\theta,\lambda}}\left[\exp\left\{itW_1-\frac{t^2}{2}V_0\right\}\left\{1-\frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}[W_0^3\mid\mathcal{F}_1]\right\}\Big|\mathcal{F}_2\right][1+O(\rho^{-1}L^{-2d+8\epsilon})]$$
$$=\exp\left\{-\frac{t^2}{2}E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\right\}\Big\{1+E^{\mu_{L,\theta,\lambda}}[B|\mathcal{F}_2]$$
$$+\frac{1}{2!}\{E^{\mu_{L,\theta,\lambda}}[B^2|\mathcal{F}_2]+E^{\mu_{L,\theta,\lambda}}\left[\frac{B^3}{3!}-\frac{it^3}{3!}W_0^3\mid\mathcal{F}_2\right]+O(\rho^{-1}L^{-2d+8\epsilon})\Big\}. \tag{9.10}$$

By definition, $E^{\mu_{L,\theta,\lambda}}[B|\mathcal{F}_2]=0$. The expectation of $B^3$ can be estimated by

$$E^{\mu_{L,\theta,\lambda}}[\frac{B^3}{3!}\mid\mathcal{F}_2]=-\frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}[W_1^3\mid\mathcal{F}_2]+O(L^{-d+\epsilon}\rho^{-1}L^{-d+\epsilon})$$

The quadratic term can be estimated by

$$E^{\mu_{L,\theta,\lambda}}[B^2|\mathcal{F}_2]=-t^2E^{\mu_{L,\theta,\lambda}}[W_1^2|\mathcal{F}_2]-\frac{it^3}{2}E^{\mu_{L,\theta,\lambda}}\left[W_1\ \{V_0-E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\ \}\mid\mathcal{F}_2\right]$$
$$+\frac{t^4}{4}E^{\mu_{L,\theta,\lambda}}\left[\left\{V_0-E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\right\}^2\Big|\mathcal{F}_2\right]+O(L^{-2d+2\epsilon}\rho^{-1})$$

Collecting the terms of order $t^3$, we have that these terms summing up to

$$-\frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}\left[W_0^3+W_1^3+3W_1\{V_0-E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\ \}\mid\mathcal{F}_2\right]=-\frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}\left[(W_0+W_1)^3\mid\mathcal{F}_2\right]$$

From the mixing condition and the assumption on $t$,

$$\frac{t^4}{2}E^{\mu_{L,\theta,\lambda}}\left[\left\{V_0-E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]\right\}^2\Big|\mathcal{F}_2\right]\le\text{const.}\rho^{-1}L^{-2d+4\epsilon}$$

Let $V_1=E^{\mu_{L,\theta,\lambda}}[W_1^2|\mathcal{F}_2]+E^{\mu_{L,\theta,\lambda}}[V_0|\mathcal{F}_2]$. We have thus proved

$$\psi_{L,\theta}^{(k)}(y)=E^{\mu_{L,\theta,\lambda}}\Big\{\exp\{i2\pi kL^{-d}\sum_{i>1}W_i\}\exp[-\frac{t^2}{2}V_1]$$
$$\times\left(1-\frac{it^3}{3!}E^{\mu_{L,\theta,\lambda}}\left[(W_0+W_1)^3\mid\mathcal{F}_2\right]+O(L^{-2d+2\epsilon}\rho^{-1})\right)\Big\}$$

We can now repeat this procedure. Since we have to repeat it $L^d$ times, the error become

$$L^dO(L^{-2d+2\epsilon}\rho^{-1})\le O(L^{-d+2\epsilon}\rho^{-1})$$

This proves Lemma 9.1.

**Appendix**

We reproduce from [SY] the proof of Lemma 4.3 in this appendix. For simplicity, we assume the Hamiltonian is given by

$$H(\eta) = -\beta \sum_{<x,y>} \eta(x)\eta(y).$$

for some $\beta > 0$. The general case can be proved in a similar way.

**Proof:** For simplicity of notation, we assume $z = (0,0), y = (2\ell, 0)$ and the Hamiltonian contains only nearest neighbor interactions. Let $A = \{(0,0), (2,0), \ldots, (2\ell, 0)\} \subset \Lambda$. Let us label these lattice sites by $j = 0, \ldots, \ell$. We condition on $\eta^c = \{\eta(x) | x \in \Lambda \setminus A\}$. Because $H$ is nearest neighbor, the conditional measure is of the form

$$(\prod_{j=0}^{\ell} p_j(\eta(j)|\eta^c))h(\sum_{j=0}^{\ell} \eta(j), \eta^c).$$

Expectations with respect to this measure are denoted by $< \cdot >_{\eta^c}$ . The function $h$ ensures the global constraint on the density and the Gibbs factor read

$$p_j(\eta(j)|\eta^c) = (\exp[\eta(j)E_j(\eta^c)] + \exp[(\eta(j) - 1)E_j(\eta^c)])^{-1},$$

where $E_j$ takes only finite number of values. Accordingly we partition $A$ into the $2d + 1$ disjoint sets $A_r = \{j \in A : E_j = \beta r\}, r = 0, 1, \ldots, 2d$. The basic idea is to perform exchanges first only within $A_0$, then within $A_1$, etc..

We start with $A_0$ and label $A_0 = \{y_j : j = 1, \ldots, n\}, y_j < y_{j+1}, |A_0| = n$. We also set $y_0 = 0$ and $y_{n+1} = \ell$, provided $y_n < \ell$. Let $T_{xy}u(\eta) = u(\eta^{xy})$. Then

$$T_{0\ell} = T_{y_0 y_1} \ldots T_{y_n y_{n+1}} T_{y_{n-1} y_n} \ldots T_{y_0 y_1}.$$

If either $y_1 = 0$ or $y_n = \ell$, then the corresponding factors are omitted. We write the telescoping sum

$$T_{0,\ell}u - u = T_{y_0 y_1} \ldots (T_{y_0 y_1}u - u) + T_{y_0 y_1} \ldots (T_{y_2 y_1}u - u) + \ldots + (T_{y_0 y_1}u - u). \quad (A.1)$$

Now $(T_{y_j y_{j+1}} f)^2 = T_{y_j y_{j+1}} f^2$ and $< T_{y_j y_{j+1}} f >_{\eta^c} = < f >_{\eta^c}$ provided $1 \le j \le n-1$. For the end points we use

$$< T_{y_0 y_1} |f| >_{\eta c} \le const. < |f| >_{\eta^c}, < T_{y_n y_{n+1}} |f| >_{\eta^c} \le e^{2d|\beta|} < |f| >_{\eta^c} .$$

Then, using (A.1) and Schwarz inequality, we arrive at

$$\ell^{-1} < (T_{0\ell} u - u)^2 >_{\eta^c} \le const. \sum_{j=0}^{n} (y_{j+1} - y_j)^{-1} < (T_{j,j+1} u - u)^2 >_{\eta^c} . \qquad (A.2)$$

If either $y_1 = 0$ or $y_n = \ell$, then the corresponding summands in (A.2) have to be omitted. Note that terms on the right hand side of (A.2) are normalized by the jump length just as on the right. Thus whenever $y_{j+1} - y_j > 1$ we may iterate our procedure for each isolated interval separately, now employing the subset $A_1$ instead of $A_0$ etc.. Then

$$\ell^{-1} < (T_{0,\ell} u - u)^2 >_{\eta^c} \le const. \sum_{j=0}^{\ell-1} < (T_{jj+1} u - u)^2 >_{\eta^c} . \qquad (A.3)$$

Average over $\eta^c$ and use that

$$< (T_{x,x+2e_1} u - u)^2 >_\Lambda \le const. [< (T_{x,x+e_1} u - u)^2 >_\Lambda + < (T_{x+e_1,x+2e_1} u - u)^2 >_\Lambda].$$

Inserting this bound in Eq. (A.3) yields Lemma 4.3.

**References**


Davies, E. B., Gross, I., and Simon, B.: Hypercontractivity: A bibliographical review, in *Ideas and Methods of Mathematics and Physics, in Memoriam of Raphael Hoegh-Krohn*, Eds., S. Albeverio, J. E. Fenstand, H. Holden, T. Lindstrom, Cambridge University Press, 1992.

[D] Davies, E. B., *Heat Kernels and Spectral Theory* Cambridge University Press, 1989.

[DS] Deuschel, J. and Stroock, D. W, *Large Deviations*, Academic Press, San Diego, 1989.

[DSa] P. Diaconis and L. Saloff-Coste: private communication.

[F] Funaki, T : private communication.

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

[GPV] M. Guo, G. C. Papanicolau and S. R. S. Varadhan, Non linear diffusion limit for a system with nearest neighbor interactions, *Comm. Math. Phys.* **118**, 31–59, 1988.

[HS] Holley, R. and Stroock, D, Logarithmic Sobolev inequalities and stochestic Ising models, *J. Stat. Phys.*, **46**, 1159-1194, 1987.

[LY] S.L. Lu and H.T. Yau: Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics, *Comm. Math. Phys.* **156**, 399–433, 1993.

[MO] Martinelli, F. and Olivieri, E.: Approach to equilibrium of Glauber dynamics in the one phase region, I and II: Commun. Math. Phys. **161**, 447-514, 1994.

[MOS] Martinelli, F., Olivieri, E. and Schomann, R.: For 2-D lattice spin systems Weak Mixing Implies Strong Mixing, to appear in Commun. Math. Phys.

[N] Newman, C.: Private communication.

[P] Petrov, V. V.: Sums of independent random variables, Springer-Verlag, Berlin, Heidelberg, New York, 1975.

[S] Stroock, D. W.: Logarithmic Sobolev Inequalities for Gibbs States, LMN, **1563**, *Dirichlet Forms*, ed Dell'Antonio, G. and Mosco, U., Springer, 1992.

[SY] Spohn, H. and Yau, H.-T.: Bulk diffusivity of lattice gases close to criticality, to appear in Jour. Stat. Phys.

[SZ] Stroock, D. and Zegarlinski, B.: The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition, Commun. Math. Phys. **144**, 303–323, 1992; see also, The logarithmic Sobolev inequality for discrete spin systems on a lattice, *Commun. Math. Phys.* 149 (1992) 175–193.

[VY] Varadhan, S.R.S. and Yau, H.-T.: Large Deviation of lattice gases with Kawasaki dynamics, in preparation.

[Y] Yau, H.-T.:Logarithmic Sobolev Inequalities for generalized simple exclusion processes, preprint.

[Z] Zegarlinski, B.: Dobrushin uniqueness theorem and logarithmic Sobolev inequalities, *J. Funct. Anal.* 105 (1992) 77–111.