Logarithmic Sobolev Inequality for Generalized Simple Exclusion Processes

Horng-Tzer Yau*

Courant Institute, New York University

New York, NY 10012

e-mail: yau@math.nyu.edu

Abstract

Let $\bar{\mu}$ be a probability measure on the set $\{0, 1, \cdots, R\}$ for some $R \in \mathbb{N}$. Let Λ_L be a cube of width L in $\mathbb{Z}^d$. Denote by $\mu^{gc}_{\Lambda_L}$ the (grand canonical) product measure with $\bar{\mu}$ as the marginal measure; here the superscript indicates the grand canonical ensemble. The canonical ensembles, denoted by $\mu^c_{\Lambda_L,n}$, are defined by conditioning $\mu^{gc}_{\Lambda_L}$ given that the total number of particle is n. Consider the zero-range exclusion dynamics where each particle performs random walk with rates depending only on the number of particles at the same site. The rates are chosen such that, for every n and L fixed, the measure $\mu^c_{\Lambda_L,n}$ is reversible. We prove the logarithmic Sobolev inequality in the sense 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 denote the Dirichlet form of the dynamics. The dependence on L is optimal.

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

1 Introduction.

Suppose that $\mathcal{L}$ is the generator of a dynamics and that μ 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 μ. A logarithmic Sobolev inequality for this system states that the entropy of a probability density f with respect to μ 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 κ^{-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 κ^{-1} provided some mild conditions on the systems are given [HS,S, SZ]. See [S, DSC] for a recent review.

In this article, we concentrate on a very simple system. Let $\bar{\mu}$ be a probability measure on the set $\{0, 1, \cdots, R\}$, for some $R \in \mathbb{N}$. Let Λ_L be a cube of width L in $\mathbb{Z}^d$. Denote a configuration by $\eta = \{\eta_x \in \{0, 1, \cdots, R\} : x \in \Lambda_L\}$. We will interpret η_x as the number of particles at site x, required to be less than or equal to R. Let $\mu^{gc}_{\Lambda_L} = \mu^{gc}_L$ be the (grand canonical) product measure with the marginal measure $\bar{\mu}$. Here the superscript, gc, indicates that the measure considered is the grand canonical ensemble in the terminology of statistical physics. The canonical ensembles, denoted by $\mu^c_{\Lambda_L,n} = \mu^c_{L,n}$, are defined by conditioning $\mu^{gc}_{\Lambda_L}$ given that the total number of particles is n. Each particle performs random walk with rates depending only on the number of particles at the same site. Since the maximum number of particles allowed at each site is R, any walk into a site with R particles is suppressed. Furthermore, the total number of particles is conserved by the dynamics because there is no creation or annihilation of particles. The rates are chosen such that, for every n and L fixed, the measure $\mu^c_{\Lambda_L,n}$ is reversible. This is the usual zero range model with the restriction that the maximum number of particles per site is R. The main result of this paper states that

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

with

$$\kappa_{L,n} \leq CL^2 \tag{1.2}$$

for some constant independent of n or L. It is easy to check that $\kappa_{L,n} \geq CL^2$ by using test functions. Hence (1.2) identifies the dependence of $\kappa_{L,n}$ on L. In the special case $n = 1$, the model reduces to the usual random walk model and the logarithmic Sobolev inequality becomes the familiar logarithmic Sobolev inequality for the discrete Laplacian.

In the special case $R = 1$, our model becomes the well-known symmetric simple exclusion process. In this case, [DSh] prove that $\kappa_{L,n} \leq CL^2 \log L$. For $R > 1$, we are not aware of nontrivial result on $\kappa_{L,n}$. The spectral gap of our model with $R = \infty$ is proved in [LSY] with some restrictions on the marginal $\bar{\mu}$.

Our proof simplifies substantially when $R = 1$ and, in particular, section 6 is not needed in this case. We shall explain this in more details in section 4; for now we observe that, in the case $R = 1$, $\mu^c_{\Lambda_L,n}$ is simply the counting measure restricted to the subspace of n particles. Hence the probability of finding any configuration η is a constant independent of η. This is similar to the Lebesgue measure in Eucildean space. The picture changes if $R > 1$. The probability of finding a configuration η depends on η. In a sense, this situation corresponds to the case $\exp[h(x)]dx$ instead of Lebesgue measure.

Recently, there are large numbers of articles concerning the logarithmic Sobolev inequality. Apart from those mentioned above, systematic studies of Glauber dynamics of lattice gases were examined by, e.g., [HS, SZ, MO, LY]. The Glauber dynamics are fundamentally different from the exclusion process considered in this article and they are not conservative in the sense that the total number of particles may change. The invariant measures for these dynamics are the grand canonical measures $\mu^{gc}_{\Lambda_L}$ instead of $\mu^c_{\Lambda_L,n}$. For non-conservative dynamics, the constant κ in (1.1) is expected to be independent of the width of the cube, L, unless phase transitions occur. A local disturbance is expected to stay local and relaxes to equilibrium exponentially fast. On the other hand, no uniform exponential relaxation is allowed for conservative dynamics because $\kappa_{L,n} \sim L^2$. 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. For the product measures considered in this article, $\langle \eta_x; \eta_y \rangle_{\mu^{gc}_{\Lambda_L}} = 0$ unless $x = y$, while $\langle \eta_x; \eta_y \rangle_{\mu^c_{L,n}} \sim -L^{-d}$. This is due in part to the following observation: if the number of particles at a site x increases by one, a particle somewhere on the lattice Λ_L needs to be removed because of the conservation of particles. Assuming the probability to remove such a particle from a site is uniform on Λ_L, we obtain that the order of magnitude is L^{-d}. This negative correlation, albeit small, is in a sense one of the the underlying reason why (1.2) holds.

Our method is based on the martingale approach of $[LY]$. We shall outline the martingale approach in section 3. The key technical input, Theorem 3.1, is then proved in section 4. Sections 5 and 6 contain proofs of lemmas assumed in section 4.

2 Statement of Main Result.

Let $\Lambda = \Lambda_L$ be a cube of width L in $\mathbf{Z}^d$ and let $X^{\Lambda_L} = \{0, 1, ..., R\}^{\Lambda_L}$ be the configuration space. Denote a typical configuration by $\eta = \{\eta_x\}$ where $\eta_x \in \{0, 1, ..., R\}$ for $x \in \Lambda_L$. Let $\bar{\mu}$ be a probability measure on the set $\{0, 1, ..., R\}$ such that $\bar{\mu}(\eta_0 = \alpha) > 0$ for any $\alpha \in \{0, 1, ..., R\}$. Let p_α be the probability of $\eta_0 = \alpha$, i.e.,

$$P^{\bar{\mu}}[\eta_0 = \alpha] := p_\alpha, \quad \alpha = 0, 1, ..., R. \tag{2.1}$$

Let $\mu^{gc}_{\Lambda_L} = \mu^{gc}_L$ be the grand canonical product with marginal $\bar{\mu}$ on Λ_L and let $\mu^c_{\Lambda_L, n} = \mu^c_{L,n}$ be the canonical measure obtained from μ^{gc}_L by conditioning on the total number of particles n.

For $z \in \mathbb{Z}^d$, let $|z| = \sum_i |z_i|$. For any function f on Λ define the Dirichlet form

$$D_\Lambda(f) = \sum_{x, x+e \in \Lambda, |e|=1} D_{x,x+e}(f),$$

$$D_{x,x+e}(f) = E^{\mu^c_{L,n}}[\{T_{x,x+e}f\}^2 + \{T_{x+e,x}f\}^2], \tag{2.2}$$

where

$$\{T_{x,y}f\}(\eta) = f(T_{x,y}\eta) - f(\eta) \tag{2.3}$$

and

$$(T_{x,y}\eta)_z := (\eta^{x,y})_z = \begin{cases} \eta_x - 1, & \text{if } z = x, \eta_x > 0 \text{ and } \eta_y < R \\ \eta_y + 1, & \text{if } z = y, \eta_x > 0 \text{ and } \eta_y < R \\ \eta_z & \text{otherwise .} \end{cases} \tag{2.4}$$

From this definition, $T_{xy}\eta$ denotes the configuration obtained by moving a particle from x to y. When $x + e$ is not in Λ, we define $D_{x,x+e}(f) = 0$ in this case.

Theorem A. *Let $u(L)$ be the smallest constant such that for any total number of particles n, any $L_i \le L$, $i = 1, \cdots, d$, and any probability density f, we have*

$$E^{\mu^c_{\Lambda_{L_1 \times \cdots \times L_d}, n}}[f \log f] \le u(L) L^2 D_{\Lambda_{L_1 \times \cdots \times L_d}, n}(\sqrt{f}); \tag{2.5}$$

here the Dirichlet form is defined by (2.2). Then there is a constant $C < \infty$ independent of L or n such that $u(L) \le C$.

3 Outline of Martingale Approach

We now prove Theorem A using the martingale decomposition as in [LY]. The key estimate is Theorem 3.1 which will be proved in sections 4-6. We will assume $d = 2$ for simplicity and that Theorem A holds for Λ of width smaller than $L \times L$. We then prove that Theorem A holds for Λ of width $L \times 2L$ and then for Λ of width $2L \times 2L$. This would then conclude Theorem A. As the proofs from $L \times L$ to $L \times 2L$ and from $L \times 2L$ to $2L \times 2L$ are identical, we will only prove Theorem A for Λ of width $L \times 2L$ assuming the statement is true for Λ of width $L \times L$. Let us order the sites in the lower half of Λ lexicographically by $i = 1, 2, \ldots, L^2$ and denote the upper half of Λ by Λ_0. More precisely, suppose the lower half of the cube is given by $\{(x,y) \in \mathbb{Z}^2 : 0 \le x \le L-1, -L+1 \le y \le 0\}$ and the upper half is given by $\{(x,y) \in \mathbb{Z}^2 : 0 \le x \le L-1, 1 \le y \le L\}$. Then according to the order, $\eta_1 = \eta_{(0,0)}, \eta_2 = \eta_{(1,0)}, \cdots \eta_L = \eta_{(L-1,0)}$ etc. Throughout section 3 to section 5, all measures involved are the canonical measures. *We shall drop all superscripts, c , in these three sections.*

Proof of Theorem A.

Step 1. Denote by $\mathcal{F}_j$ the σ-algebra generated by $\eta_j, \eta_{j+1}, \cdots$. Define the marginal probability density f_j by $f_0 = f$,

$$f_j(\eta_j, \eta_{j+1} \ldots) = E^{\mu_{L\times 2L,n}}[f|\mathcal{F}_j],\ 1 \le j \le L^2, \tag{3.1}$$

and $f_j = 1$ for $j \ge L^2 + 1$. 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 $j = L^2$. By definition of u, for any cube Λ of width $L_1 \times L_2$ with $L_i \le L$,

$$E^{\mu_{\Lambda,n}}[f \log f] \le u(L) L^2 D_\Lambda(\sqrt{f}). \tag{3.3}$$

Since Λ_0 is a cube of width $L \times L$,

$$\begin{aligned} &E^{\mu_{L\times 2L,n}}[(f/f_1)\log(f/f_1)|\mathcal{F}_1] \\ &\le u(L)L^2 \sum_{x,x+e\in\Lambda_0,|e|=1} E^{\mu_{L\times 2L,n}}[\{T_{x,x+e}\sqrt{(f/f_1)}\}^2 + \{T_{x+e,x}\sqrt{(f/f_1)}\}^2|\mathcal{F}_1]. \end{aligned}$$

Note that for $x, x+e \in \Lambda_0$,

$$T_{x,x+e}\sqrt{(f/f_1)} = (T_{x,x+e}\sqrt{f})/\sqrt{f}_1$$

For simplicity, we denote

$$E^{\mu_{L\times 2L,n}}[\{T_{x,x+e}\sqrt{f}\}^2 + \{T_{x+e,x}\sqrt{f}\}^2] = D_{x,x+e}$$

Multiplying both sides by f_1 and taking expectation w.r.t. $\mu_{L\times 2L,n}$, we have

$$E^{\mu_{L\times 2L,n}}\{E^{\mu_{L\times 2L,n}}[f\log(f/f_1)|\mathcal{F}_1]\} \le u(L)L^2 \sum_{x,x+e\in\Lambda_0,|e|=1} D_{x,x+e}(\sqrt{f}). \tag{3.4}$$

Step 2. We now bound each summand on the right side of (3.2). The result can be stated as the following Theorem 3.1 proved in next section.

Theorem 3.1 *Suppose (3.3) holds for Λ of width $L_1 \times L_2$ with $L_i \le 2L$. Let $\bar{f}_z$ be the marginal of f at z. Then for any $z \in \Lambda$ and $\epsilon_0 > 0$ fixed there is a constant C_1 depending on ϵ_0 but independent of L or z such that*

$$E^{\mu_{\Lambda,n}}[\bar{f}_z \log \bar{f}_z] \le C_1 L Av_{y\in\Lambda} D^{\gamma_{zy}}(\sqrt{f}) + C_1 L^{2-d} D_\Lambda(\sqrt{f}) + \epsilon_0 u(2L) L^{2-d} D_\Lambda(\sqrt{f}) \tag{3.5}$$

where Av denotes average. Here γ_{zy} is a path from z to y chosen in a fixed way (e.g., from $z = (z_1, z_2)$ to (z_1, y_2) via a straight line and then to $(y_1, y_2) = y$ via a straight line) and $D^{\gamma_{z,y}}(\sqrt{f})$ is defined by

$$D^\gamma(\sqrt{f}) = \sum_{i=1}^{|\gamma|} D_{\gamma(i-1),\gamma(i)}(\sqrt{f}), \tag{3.6}$$

where $\gamma(i)$ is the i-th point in the path γ and $|\gamma|$ is the length of the path.

Now applying Theorem 3.1 to bound (3.2) for $j > 0$, we have

$$\begin{aligned}
&\sum_{j=1}^{\infty} E^{\mu_{L\times 2L,n}} E^{\mu_{L\times 2L,n}}[f_j \log(f_j/f_{j+1})|\mathcal{F}_{j+1}] \\
\le\ & C_1 L^{1-d} \sum_{x\in\Lambda_{L\times 2L}} \sum_{y\in\Lambda_{L\times 2L}} D^{\gamma_{xy}}(\sqrt{f}) + C_1 L^{2-d} D_{\Lambda_{L\times 2L}}(\sqrt{f}) + \epsilon_0 u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) \\
\le\ & C_2 L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) + \epsilon_0 u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}).
\end{aligned} \tag{3.7}$$

Strictly speaking, we have applied Theorem 3.1 to domains slightly different from a cube due to the conditioning on $\mathcal{F}_{j+1}$. To avoid this difficulty, one has to work with "generalized cubes" as in [LY]. One can also check the proof of Theorem 3.1 applied to "generalized cubes" as well. As this has been done in [LY], we omit the details.

Step 3. Combining (3.2), (3.4) and (3.7), one has

$$\begin{aligned}
&E^{\mu_{L\times 2L,n}}[f \log f] \\
&\le C_2 L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) + \epsilon_0 u(2L) L^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}) \\
&+ u(L) L^2 \sum_{x,x+e\in\Lambda_0} \sum_{|e|=1} D_{x,x+e}(\sqrt{f}).
\end{aligned}$$

Switching the role of Λ_0 and $\Lambda \backslash \Lambda_0$ and taking the average, we have

$$E^{\mu_{L\times 2L,n}}[f \log f] \le [C_2 + u(L)/8 + \epsilon_0 u(2L)](2L)^2 D_{\Lambda_{L\times 2L}}(\sqrt{f}). \tag{3.8}$$

Note that Λ in (3.8) is of width $L \times 2L$. We can repeat steps 1-3 once more and obtain for Λ of width $2L \times 2L$

$$E^{\mu_{2L\times 2L,n}}[f \log f] \le [2C_2 + u(L)/2 + 2\epsilon_0 u(2L)](2L)^2 D_{\Lambda_{2L\times 2L}}(\sqrt{f}). \tag{3.9}$$

From the definition of $u(2L)$, we have thus proved that

$$u(2L) \leq [2C_2 + u(L)/2 + 2\epsilon_0 u(2L)]$$

Since ϵ_0 can be chosen as small as we wish, say $\epsilon_0 = 1/16$, we have proved that,

$$u(2L) \leq C + 2u(L)/3.$$

From the induction or the Gronwal type inequality we have proved Theorem A assuming Theorem 3.1.

4 Proof of Theorem 3.1.

We now prove Theorem 3.1.

Step 1. The first step in proving Theorem 3.1 is the following lemma 4.1 which provides a bound on the logarithmic Sobolev constant for probability density on one site.

Lemma 4.1 *Let ν be a probability measure on the set $S = \{0, 1, 2, \ldots, R\}$ with $P^\nu[\eta = \alpha] := q_\alpha > 0$. Suppose there is a constant $\gamma > 0$ independent of ρ such that*

$$(i) \quad \rho := E^\nu[\eta] \leq R/2\,, \qquad (ii) \quad q_\alpha \leq \gamma \rho q_{\alpha-1},\ 1 \leq \alpha \leq R. \tag{4.1}$$

Let f be a probability density relative to ν, namely, $\sum_{\alpha=0}^R f(\alpha) q_\alpha = 1$. Then

$$s(f/\nu) = \sum_{\alpha=0}^R q_\alpha f(\alpha) \log f(\alpha) \leq C_1\, \rho |\log \rho| D_\nu(\sqrt{f}) \tag{4.2}$$

where D_ν is the Dirichlet form defined by

$$D_\nu[g] = \sum_{\alpha=0}^R (g(\alpha+1) - g(\alpha))^2 q_\alpha \tag{4.3}$$

with the convention $g(R) = g(R+1)$. The constant C_1 in (4.2) depends on q_α only through γ but is independent of ρ or f.

Furthermore, for a constant $\kappa > 1$ let $A = \{\alpha | f(\alpha) > \kappa f(0)\}$ and $B = \{0, 1, \ldots, R\} \backslash A$ and decompose the entropy into $s(f/\nu) = s_A(f/\nu) + s_B(f/\nu)$ with

$$s_A(f/\nu) = \sum_{\alpha \in A} f(\alpha) \log f(\alpha) q_\alpha \tag{4.4}$$

and $s_B = s - s_A$. Then the factor $\log \rho$ can be omitted if restricted to s_B, namely,

$$s_B(f/\nu) \leq C_2 \rho D(\sqrt{f}) \tag{4.5}$$

with the constant C_2 in (4.5) depending on κ and γ but is independent of ρ or f.

Remark. The existence of a constant such that (4.2) holds is obvious since the state space S is finite. The main point of this Lemma is the identification of the dependence of ρ. The proof of Lemma 4.1 will be given in section 5.

Return to the proof Theorem 3.1. We assume that without loss of generality, $\rho = n/|\Lambda| \le R/2$. The case $\rho \ge R/2$ can be handled similarly. Recall $\bar{f}_z$ is the marginal of f at z. Apply Lemma 4.1 to the marginal distribution of $\mu_{\Lambda,n}$ at z,

$$E^{\mu_{\Lambda,n}}[\bar{f}_z \log \bar{f}_z] \le C\rho|\log\rho| \sum_{\alpha=0}^{R-1} \{\sqrt{\bar{f}_z(\alpha+1)} - \sqrt{\bar{f}_z(\alpha)}\}^2 q_\alpha \tag{4.6}$$

where $q_\alpha = E^{\mu_{\Lambda,n}}[\eta_z = \alpha]$. The assumption (4.1) can be checked easily with γ depends only only on p_α (2.1) in the definition of the marginal measure $\bar{\mu}$.

Step 2. The right side of (4.6) can be bounded using the following lemma which will be proved in section 5. *For the rest of this paper, we shall use C to denote any constant independent of L, ρ or f and depending only on p_α (2.1) in the definition of the marginal measure $\bar{\mu}$.*

Lemma 4.2. *Suppose g is a nonnegative function on the configuration space of $\Lambda = \Lambda_L$ and $0 \notin \Lambda$. Define $\bar{g}(m) = E^{\mu_{\Lambda,m}}[g]$. Then*

$$\{\sqrt{\bar{g}(m+1)} - \sqrt{\bar{g}(m)}\}^2 \le A_1 + A_2$$

where

$$\begin{aligned} A_1 &= C(R - m/L^d)^{-1} G(m+1) Av_x E^{\mu_{\Lambda,m+1}}[\{\sqrt{g(T_{x0}\eta)} - \sqrt{g(\eta)}\}^2] \\ A_2 &= C(R - m/L^d)^{-2} E^{\mu_{\Lambda,m}}[g; \tilde{h}]^2 \{\bar{g}(m+1) + \bar{g}(m)\}^{-1}, \end{aligned} \tag{4.7}$$

where $G(m+1) = L^d(m+1)^{-1}, \tilde{h} = Av_{x\in\Lambda} h(\eta_x)$ with

$$h(a) = 1_{a<R}\, p_a^{-1}\, p_{a+1}(a+1) \tag{4.8}$$

and p_a is defined in (2.1). Furthermore, the constant in (4.7) depends on p_α but is independent of m/L^d.

Remark. In the case $R = 1$, we can assume $p_0 = p_1 = 1/2$, because we are only interested in the measure $\mu_{\Lambda,m}$. Hence $\tilde{h} = Av_{x\in\Lambda}(1 - \eta_x) = 1 - m/L^d$ is a constant and $A_2 = 0$. Therefore, the following Lemma 4.4 and its proof, section 6, will not be needed in this case.

Applying Lemma 4.2 with $\Lambda \to \Lambda \setminus \{z\}, \mu_{\Lambda_L,m} \to \mu_{\Lambda,n}|_{\eta_z=n-m}$, $0 \to z$ and

$$\bar{g}(m) \to E^{\mu_{\Lambda,n}}[f|\eta_z = n-m] = \bar{f}_z(n-m) = \bar{f}_z(\alpha+1),$$

one has

$$\left\{\sqrt{\bar{f}_z(\alpha+1)} - \sqrt{\bar{f}_z(\alpha)}\right\}^{-2} \le B_1(\alpha) + B_2(\alpha)$$

where

$$
\begin{aligned}
B_1(\alpha) &= C[R-(n-\alpha-1)/L^d]^{-1}G(n-\alpha)Av_x E^{\mu_{\Lambda,n}}\Big[\Big\{\sqrt{f(T_{xz}\eta)}-\sqrt{f(\eta)}\Big\}^2\Big|\eta_z=\alpha\Big] \\
B_2(\alpha) &= C[R-(n-\alpha-1)/L^d]^{-2}E^{\mu_{\Lambda,n}}\Big[f;\tilde{h}\Big|\eta_z=\alpha+1\Big]^2\Big[\bar{f}_z(\alpha+1)+\bar{f}_z(\alpha)\Big]^{-1}
\end{aligned}
$$

Define

$$
\begin{aligned}
&\Omega_1 = C(L^d/n)Av_x E^{\mu_{\Lambda,n}}\Big[\Big\{\sqrt{f(T_{xz}\eta)}-\sqrt{f(\eta)}\Big\}^2\Big] \\
&\Omega_2 = C(L^d/n)E^{\mu_{\Lambda,n}}\Big(E^{\mu_{\Lambda,n}}[f;\tilde{h}|\eta_z]^2\bar{f}_z(\eta_z)^{-1}\Big),
\end{aligned}
\tag{4.9}
$$

Recall $q_\alpha = E^{\mu_{\Lambda,n}}[\eta_z=\alpha]$. We claim that

$$
\sum_{\alpha=0}^{R-1}\Big\{\sqrt{\bar{f}_z(\alpha+1)}-\sqrt{\bar{f}_z(\alpha)}\Big\}^2 q_\alpha \le \Omega_1+\Omega_2 \tag{4.10}
$$

Since $n/L^d \le R/2$, the prefactor $[R-(n-\alpha-1)/L^d]^{-1}$ in B_1 and B_2 is bounded above. From Lemma 4.2, $G(n-\alpha)\le CL^d/n$. Multiplying both sides of $B_1(\alpha)$ by q_α and summing over α , we have

$$
\sum_{\alpha=0}^{R-1} B_1(\alpha)q_\alpha \le \Omega_1.
$$

Multiplying both sides of $B_2(\alpha)$ by q_α and summing over α , we have

$$
\sum_{\alpha=0}^{R-1} B_1(\alpha)q_\alpha \le C(n/L^d)\sum_\alpha q_\alpha E^{\mu_{\Lambda,n}}[f;\tilde{h}|\eta_z=\alpha+1]^2 k(\alpha+1)
$$

where

$$
k(\alpha+1) = \Big[\bar{f}_z(\alpha+1)+\bar{f}_z(\alpha)\Big]^{-1} \le [\bar{f}_z(\alpha+1)]^{-1}.
$$

By definition of q_α, there is a constant γ depending only on p_α (2.1) such that

$$
\gamma^{-1}q_\alpha(n/L^d) \le q_{\alpha+1} \le \gamma q_\alpha(n/L^d). \tag{4.11}
$$

Hence for any function k and g,

$$
\begin{aligned}
\sum_\alpha q_\alpha E^{\mu_{\Lambda,n}}[g|\eta_z=\alpha+1]^2k(\alpha+1) &\le C(L^d/n)\sum_\alpha q_{\alpha+1}E^{\mu_{\Lambda,n}}[g|\eta_z=\alpha+1]^2k(\alpha+1) \\
&= C(L^d/n)E^{\mu_{\Lambda,n}}\Big(E^{\mu_{\Lambda,n}}[g|\eta_z]^2k(\eta_z)\Big)
\end{aligned}
\tag{4.12}
$$

This proves the bound on B_2 and concludes (4.10).

Step 3. The following two lemmas give bounds on Ω_1 and Ω_2. Lemma 4.3 will be proved in section 5 ; Lemma 4.4 will be proved in section 6.

Lemma 4.3 *Suppose γ_{xy} is a path from x to y with $|\gamma_{xy}| \le CL$. Then*

$$E^{\mu_{\Lambda,n}}[\{\sqrt{f(T_{xy}\eta)} - \sqrt{f(\eta)}\}^2] \le CLD^{\gamma_{xy}}(\sqrt{f}).$$

where $D^{\gamma_{xy}}$ is defined in (3.6).

Lemma 4.4 *Suppose F is a bounded function and Λ is of width smaller than $2L \times 2L$. Then for any $\epsilon_0 > 0$ and any probability density f,*

$$\langle f; Av_x F(\eta_x)\rangle^2_{\mu_{\Lambda,n}} \le [C + \epsilon_0 u(2L)]L^2 Av_{x,x+e\in\Lambda,|e|=1} D_{x,x+e}(\sqrt{f}) \tag{4.13}$$

Here $\langle f; g\rangle = \langle fg\rangle - \langle f\rangle\langle g\rangle$ and the constant depends on ϵ_0 and F.

Recall we have assumed that $\rho = n/L^d \le R/2$. From Lemma 4.3 we can bounded Ω_1 by

$$\Omega_1 \le C(L^d/n)LAv_y D^{\gamma_{zy}}(\sqrt{f})$$

Applying Lemma 4.4 to the measure $\mu_{\Lambda,n}|_{\eta_z}$ and then taking expectation w.r.t $\mu_{\Lambda,n}$, we can bound Ω_2 by

$$\Omega_2 \le C(L^d/n)[1 + \epsilon_0 u(2L)]L^{2-d} D_\Lambda(\sqrt{f})$$

Hence if $\rho = n/L^d \le R/2$, from (4.10),

$$\sum_{\alpha=0}^{R} \{\sqrt{\bar{f}_z(\alpha+1)} - \sqrt{\bar{f}_z(\alpha)}\}^2 q_\alpha \le C(L^d/n)Q_L, \tag{4.14}$$

where

$$Q_L = C\Big\{LAv_y D^{\gamma_{zy}}(\sqrt{f}) + [1 + \epsilon_0 u(2L)]L^{2-d} D_\Lambda(\sqrt{f})\Big\}. \tag{4.15}$$

Together with (4.6), we have thus proved

$$E^{\mu_{\Lambda,n}}[\bar{f}_z \log \bar{f}_z] \le C|\log\rho| Q_L. \tag{4.16}$$

where the path γ_{xy} is the canonical path chosen in Theorem 3.1. Note that the factor L^d/n in (4.14) is canceled by ρ from (4.6). This concludes Theorem 4.1 except there is the logarithmic factor as $\rho \to 0$. This factor certainly can be bounded by $\log L$. Using this bound, one has a logarithmic Sobolev inequality with an extra $\log L$ factor. Finally we have to remove the logarithmic factor in (4.16).

Step 4. We now remove the logarithmic factor in (4.16). To remove it, it suffices to consider $\rho \ll 1$. The argument in this step will not be used for the rest of this paper. From (4.14) and the elementary inequality $a^2 \le 2b^2 + 2(a-b)^2$,

$$\bar{f}_z(\alpha+1) \le 2\bar{f}_z(\alpha) + Cq_\alpha^{-1}\rho^{-1}Q_L.$$

Hence from the definition $\bar{f}_z(\alpha) = E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}]\,\{E^{\mu_{\Lambda,n}}[1_{\eta_z=\alpha}]\}^{-1}$ and (4.11), one has for $\alpha \geq 0$ and $\rho \leq R/2$,

$$E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha+1}] \leq C\rho E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}] + CQ_L.$$

By induction,

$$E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}] \leq C[\rho^\alpha \bar{f}_z(0) + Q_L] \tag{4.17}$$

for some constant C independent of ρ, f or L.

Let Λ_ℓ be a cube of width ℓ containing z. Denote the density in Λ_ℓ by $\bar{\eta}_\ell$. Clearly, $\bar{\eta}_\ell \leq \rho L^d/\ell^d$. Let $\mathcal{F}_{\Lambda\setminus\Lambda_\ell}$ be the $\sigma-$algebra of the configurations in $\Lambda \setminus \Lambda_\ell$ and let $f_{\mathcal{F}_{\Lambda\setminus\Lambda_\ell}}$ be the conditional probability given $\mathcal{F}_{\Lambda\setminus\Lambda_\ell}$. Define $\bar{f}_{z,\mathcal{F}_{\Lambda\setminus\Lambda_\ell}}$ to be the marginal of $f_{\mathcal{F}_{\Lambda\setminus\Lambda_\ell}}$ on $\{z\}$ and define Q_ℓ as in (4.15) except L is replaced by ℓ and the Dirichlet form is defined with respect to $\mu_{\Lambda,n}|_{\mathcal{F}_{\Lambda\setminus\Lambda_\ell}}$. Note that $\bar{\eta}_\ell$ is measurable w.r.t $\mathcal{F}_{\Lambda\setminus\Lambda_\ell}$ because the total number of particles in Λ is fixed. Applying (4.17) to this setting, we have for $\rho L^d/\ell^d \leq R/2$,

$$\begin{aligned} E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}|\mathcal{F}_{\Lambda\setminus\Lambda_\ell}] &\leq C[(\bar{\eta}_\ell)^\alpha \bar{f}_{z,\mathcal{F}_{\Lambda\setminus\Lambda_\ell}}(\eta_z=0) + Q_\ell] \\ &\leq C[\rho^\alpha (L/\ell)^{\alpha d} \bar{f}_{z,\mathcal{F}_{\Lambda\setminus\Lambda_\ell}}(\eta_z=0) + Q_\ell]. \end{aligned}$$

By definition of Q_ℓ and Q_L (4.15), $E^{\mu_{\Lambda,n}}[Q_\ell] \leq C(\ell/L)Q_L$. Taking expectation with respect to $\mu_{\Lambda,n}$, we have

$$E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}] \leq C\rho^\alpha \bar{f}_z(0)\Big[(L/\ell)^{\alpha d} + (\ell/L)\Big\{\,Q_L/(\bar{f}_z(0)\rho^\alpha)\,\Big\}\Big] \tag{4.18}$$

The optimal choice of ℓ is given by

$$L/\ell = \Big\{\,Q_L/(\bar{f}_z(0)\rho^\alpha)\,\Big\}^{\frac{1}{1+d\alpha}}.$$

Hence we have

$$\bar{f}_z(\alpha) \leq C\rho^{-\alpha} E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}] \leq C\bar{f}_z(0)\Big\{\,Q_L/(\bar{f}_z(0)\rho^\alpha)\,\Big\}^{\frac{d\alpha}{d\alpha+1}} \tag{4.19}$$

provided

$$1 \leq \Big\{\,Q_L/(\bar{f}_z(0)\rho^\alpha)\,\Big\} \leq [R/(2\rho)]^{(1+d\alpha)/d}. \tag{4.20}$$

The first restriction comes from $L \geq \ell$; the second one comes from $\rho L^d/\ell^d \leq R/2$.

Suppose $Q_L[\bar{f}_z(0)\rho^\alpha]^{-1} < 1$. From (4.17)

$$\bar{f}_z(\alpha) \leq C\bar{f}_z(0)[1 + Q_L\{\bar{f}_z(0)\rho^\alpha\}^{-1}] \leq 2C\bar{f}_z(0).$$

with the constant C given by (4.17). Let

$$U = \Big\{\alpha : \bar{f}_z(\alpha) \leq 2C\bar{f}_z(0)\Big\}, \tag{4.21}$$

$$V = \left\{ \alpha : Q_L/[\bar{f}_z(0)\rho^\alpha] \geq [R/(2\rho)]^{(1+d\alpha)/d} \right\} \setminus U \tag{4.22}$$

so that (4.20) holds for $\alpha \notin U \cup V$.

Case 1: $\alpha \in U$. We can apply (4.5) on U to have

$$\sum_{\alpha:\alpha \in U} \bar{f}_z(\alpha) \log \bar{f}_z(\alpha) q_\alpha \leq C Q_L. \tag{4.22}$$

From the assumption $\rho \ll 1$, we have $q_\alpha \sim \rho^\alpha$ and thus $\bar{f}_z(\alpha) \leq C\rho^{-\alpha}$ because f is a probability density. In particular, $\bar{f}_z(0)$ is bounded above independent of ρ.

Case 2 : $\alpha \in V$. Instead of the optimal choice of ℓ, we choose $L/\ell = [R/(2\rho)]^{1/d}$. Hence from (4,18), one has

$$\bar{f}_z(\alpha) \leq C\rho^{-\alpha} E^{\mu_{\Lambda,n}}[f 1_{\eta_z=\alpha}] \leq C\bar{f}_z(0)\rho^{-\alpha} + \rho^{1/d-\alpha} Q_L$$

Since $\alpha \in V$, we have $Q_L \geq C\rho^{-1/d}$. Hence

$$\bar{f}(\alpha) \leq C\rho^{-\alpha+(1/d)} Q_L \ .$$

Together with the definition of entropy and $\log \bar{f}_z(\alpha) \leq C|\log\rho|$, we have

$$\sum_{\alpha \in V} \bar{f}_z(\alpha) \log \bar{f}_z(\alpha) q_\alpha \leq C \sum_\alpha \rho^{-\alpha+(1/d)} |\log\rho| \rho^\alpha Q_L \leq C Q_L.$$

Case 3: $\alpha \notin U \cup V$. Since (4.20) is satisfied outside $U \cup V$, (4.19) is satisfied for $\alpha \notin U \cup V$. We assume $f_z(\alpha) \geq 1$ since otherwise $\bar{f}_z(\alpha) \log \bar{f}_z(\alpha) < 0$ and we can drop this term. Since $x \log x \leq C x^\gamma$ for $x \geq 1$ and any $\gamma > 1$, let $\gamma = \frac{d\alpha+1}{d\alpha}$ we have

$$\bar{f}_z(\alpha) \log \bar{f}_z(\alpha) q_\alpha \leq C q_\alpha \left\{ \bar{f}_z(0) \left\{ Q_L/(\bar{f}_z(0)\rho^\alpha) \right\}^{\frac{d\alpha}{1+d\alpha}} \right\}^{\frac{d\alpha+1}{d\alpha}} \leq C Q_L,$$

where we have used $q_\alpha \leq C\rho^\alpha$ and $\bar{f}_z(0)$ is bounded above (since $\bar{f}_z(0) q_0 \leq 1$ and $q_\alpha \sim \rho^\alpha$).

Combining these bounds, we have thus proved that

$$\sum_\alpha \bar{f}_z(\alpha) \log \bar{f}_z(\alpha) q_\alpha \leq C Q_L.$$

This concludes Theorem 3.1.

5 Proof of Lemmas 4.1-4.3

We now prove Lemmas 4.1-4.3.

Proof of Lemma 4.1. We shall assume that $\kappa = 2$ for simplicity of notation.

Step 1. We denote all constants depending on γ but independent of ρ in the following proof simply by C. From the assumption (4.1) we have $q_0 \geq C$ and $|\log q_\alpha| \leq C \log \rho|$ with the constant independent of ρ. Also $f(\alpha) q_\alpha \leq 1$ and thus $\log f(\alpha) \leq C|\log \rho|$. Therefore,

$$s_A(f/\mu) = \sum_{\alpha \in A} q_\alpha f(\alpha) \log f(\alpha) \leq C|\log \rho| \sum_{\alpha \in A} f_\alpha q_\alpha. \tag{5.1}$$

where s_A is defined in (4.4). From the Schwarz inequality

$$[\sqrt{f(\alpha)} - \sqrt{f(0)}]^2 \leq \alpha \sum_{\delta=1}^{\alpha} (\sqrt{f(\delta)} - \sqrt{f(\delta-1)})^2 \leq R \sum_{\delta=1}^{\alpha} (\sqrt{f(\delta)} - \sqrt{f(\delta-1)})^2$$

By (ii) of (4.1), $q_\alpha \geq C q_\delta$ if $\alpha \leq \delta$. Hence from the definition of the Dirichlet form one has for $\alpha \geq 0$,

$$\begin{aligned} D(\sqrt{f}) &= \sum_{\alpha=1}^{R} [\sqrt{f(\alpha)} - \sqrt{f(\alpha-1)}]^2 q_{\alpha-1} \\ &\geq R^{-1} \sum_{\alpha=1}^{R} \sum_{\delta=1}^{\alpha} (\sqrt{f(\delta)} - \sqrt{f(\delta-1)})^2 q_{\delta-1} \\ &\geq C R^{-2} \sum_{\alpha=1}^{R} [\sqrt{f(\alpha)} - \sqrt{f(0)}]^2 q_{\alpha-1}. \end{aligned} \tag{5.2}$$

From the definition of the set A, $[\sqrt{f}(\alpha) - \sqrt{f}(0)]^2 \geq \sqrt{f}(\alpha)/16$. From (4.1) and (5.1, 5.2),

$$D(\sqrt{f}) \geq C R^{-2} \rho^{-1} \quad \sum_{\alpha \in A} f(\alpha) q_\alpha \geq C[\rho |\log \rho|]^{-1} s_A(f/\nu). \tag{5.3}$$

Step 2 It remains to bound s_B by $s_B(f/\nu) \leq C \rho D(\sqrt{f})$. Note that the logarithmic factor, $|\log \rho|$, is not needed here. From the concavity of log, one has $\log x \leq x - 1$ for all $x > 0$. Apply this bound and using $\sum_{\alpha \in S} f(\alpha) q_\alpha = 1$,

$$\begin{aligned} s_B(f) &= \sum_{\beta \in B} f(\beta) q_\beta \log f(\beta) \\ &\leq \sum_{\beta \in B} f(\beta) q_\beta \Big\{ f(\beta) - [\sum_{\alpha \in S} f(\alpha) q_\alpha] \Big\} \\ &= \sum_{\beta \in B} \sum_{\alpha \in B} q_\beta f(\beta) [f(\beta) - f(\alpha)] q_\alpha + \sum_{\beta \in B} \sum_{\alpha \in A} q_\beta f(\beta) [f(\beta) - f(\alpha)] q_\alpha \end{aligned}$$

Since $f(\beta) \leq f(\alpha)$ for $\alpha \in A$ and $\beta \in B$, the last term is negative. Hence

$$s_B(f) \leq \sum_{\beta \in B} \sum_{\alpha \in B} q_\beta f(\beta) [f(\beta) - f(\alpha)] q_\alpha = \sum_{\beta < \alpha; \alpha, \beta \in B} [f(\alpha) - f(\beta)]^2 q_\alpha q_\beta. \tag{5.4}$$

Note that (5.2) holds with 0 replaced by β and α restricted to $\alpha > \beta$. Thus multiply both sides of (5.2) (with 0 replaced by β) by q_β and sum over β,

$$D(\sqrt{f}) \geq C \sum_{\beta < \alpha; \alpha, \beta \in B} [\sqrt{f(\alpha)} - \sqrt{f(\beta)}]^2 q_{\alpha-1} q_\beta.$$

To conclude Lemma 4.1, it suffices to prove that

$$\sum_{\beta<\alpha;\alpha,\beta\in B}[f(\alpha)-f(\beta)]^2 q_\alpha q_\beta \le C\rho \sum_{\beta<\alpha;\alpha,\beta\in B}[\sqrt{f(\alpha)}-\sqrt{f(\beta)}]^2 q_{\alpha-1}q_\beta, \tag{5.5}$$

From the definition of the set B and (4.1),

$$[\sqrt{f(\alpha)}+\sqrt{f(\beta)}]^2 \le Cf(0) \le Cq_0^{-1} \le \text{ const.}$$

for any $\alpha,\beta\in B$. This, together with $q_\alpha \le Cq_{\alpha-1}\rho$, proves (5.5) and concludes lemma 4.1.

Proof of Lemma 4.2.

Step 1 Recall $h(\eta_y)$ is defined in (4.8). From the definition of covariance,

$$\bar g(m) = E^{\mu_{\Lambda,m}}[g] = E^{\mu_{\Lambda,m}}[\tilde h]^{-1}\Big\{-E^{\mu_{\Lambda,m}}[g;\tilde h] + E^{\mu_{\Lambda,m}}[g\tilde h]\Big\} \tag{5.6}$$

For each $x\in\Lambda$ fixed, change the variable to $\eta = T_{x0}\xi$ where T_{x0} is defined in (2.4). Hence

$$E^{\mu_{\Lambda,m}}[g\tilde h] = \tilde G(m) Av_{x\in\Lambda} E^{\mu_{\Lambda,m+1}}[g(T_{x0}\xi)\xi_x],$$

where $\tilde G(m)$ is some function of m independent of g. Combining these two identities one has

$$\bar g(m) = -E^{\mu_{\Lambda,m}}[\tilde h]^{-1}E^{\mu_{\Lambda,m}}[g;\tilde h] + G(m+1)Av_{x\in\Lambda}E^{\mu_{\Lambda,m+1}}[g(T_{x0}\xi)\xi_x]. \tag{5.7}$$

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

$$E^{\omega_{m+1}}[t] = G(m+1)Av_{x\in\Lambda}E^{\mu_{\Lambda,m+1}}[t_x(\eta)\eta_x]. \tag{5.8}$$

In particular, if t is independent of x then $E^{\omega_{m+1}}[t] = E^{\mu_{\Lambda,m+1}}[t]$. Hence $E^{\omega_{m+1}}[g(\eta)] = \bar g(m+1)$. Let $g(T_{*0}\eta)-g(\eta)$ denote the family of functions $(g(T_{x0}\eta)-g(\eta))_{x\in\Lambda}$. Using this identity in (5.7), we have

$$\bar g(m+1) - \bar g(m) = -E^{\mu_{\Lambda,m}}[\tilde h]^{-1}E^{\mu_{\Lambda,m}}[g;\tilde h] - E^{\omega_{m+1}}[(g(T_{*0}\eta)-g(\eta))]. \tag{5.9}$$

Step 2. From the elementary inequality

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

one has from (5.9),

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

where

$$\begin{aligned} A_3 &= 4\left\{E^{\omega_{m+1}}[g(T_{*0}\eta)-g(\eta)]\right\}^2/[\bar g(m+1)+\bar g(m)] \\ A_4 &= 4E^{\mu_{\Lambda,m}}[g;\tilde h]^2 E^{\mu_{\Lambda,m}}[\tilde h]^{-2}[\bar g(m+1)+\bar g(m)]^{-1}. \end{aligned} \tag{5.10}$$

From the Schwarz inequality

$$\begin{aligned}
&\{E^{\omega_{m+1}}[g(T_{*0}\eta)-g(\eta)]\}^2 \\
&\le E^{\omega_{m+1}}\Big[\{\sqrt{g(T_{*0}\eta)}-\sqrt{g(m)}\}^2\Big]E^{\omega_{m+1}}\Big[\{\sqrt{g(T_{*0}\eta)}+\sqrt{g(\eta)}\}^2\Big] \\
&\le 2E^{\omega_{m+1}}\Big[\{\sqrt{g(T_{*0}\eta)}-\sqrt{g(m)}\}^2\Big]E^{\omega_{m+1}}\Big[g(T_{*0}\eta)+g(\eta)\Big]
\end{aligned}$$

By definition, $E^{\omega_{m+1}}[g(\eta)] = \bar{g}(m+1)$. Also, from (5.7), one has

$$E^{\omega_{m+1}}[g(T_{*0}\eta)] = E^{\mu_{\Lambda,m}}[\tilde{h}]^{-1}E^{\mu_{\Lambda,m}}[g;\tilde{h}]+\bar{g}(m).$$

By definition, $\tilde{h}$ is bounded above and $E^{\mu_{\Lambda,m}}[\tilde{h}]^{-1}$ is bounded above by

$$E^{\mu_{\Lambda,m}}[\tilde{h}]^{-1} \le CE^{\mu_{\Lambda,m}}[Av_x 1(\eta_x < R)]^{-1} \le C(R-m/L^d)^{-1}.$$

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

$$E^{\omega_{m+1}}[g(T_{*0}\eta)] \le C\bar{g}(m)(R-m/L^d)$$

Hence

$$E^{\omega_{m+1}}\Big[g(T_{*0}\eta)+g(\eta)\Big][\bar{g}(m+1)+\bar{g}(m)]^{-1} \le C(R-m/L^d)^{-1}$$

Hence we have proved $A_3 \le A_1$ with A_1 defined in (4.7). Using again the bound on $E^{\mu_{\Lambda,m}}[\tilde{h}]^{-1}$, we have $A_4 \le A_2$. This proves Lemma 4.2.

Lemma 4.3 is indeed a statement of the Schwarz inequality. It can be proved by moving particles from x to y and followed by the Schwarz inequality as in [KOV]. The only complication comes from the number of particles per site can now be bigger than one. Let us now sketch some modifications needed.

Define the new Dirichlet form

$$\tilde{D}_\Lambda(f) = \sum_{|x-y|=1} E^{\mu_{\Lambda,n}}[(\tilde{T}_{x,y}f)^2],$$

where $\tilde{T}_{x,y}f = f(\tilde{T}_{x,y}\eta)-f(\eta)$ and

$$(\tilde{T}_{x,y}\eta)_z = \eta_y, \text{ if } z=x \quad ; = \eta_x, \text{ if } z=y \quad ; = \eta_z, \text{ otherwise.} \tag{5.11}$$

The dynamics described by (5.11) is the familiar process which exchanges configurations of nearest-neighbored sites. Clearly, the new Dirichlet form $\tilde{D}$ can be bounded by the old Dirichlet form with at most a constant factor, namely $\tilde{D} \le CD$. With the new Dirichlet form we have the following version of Lemma 4.3.

Lemma 5.1 *Suppose γ_{xy} is a path from x to y with $|\gamma_{xy}| \le CL$. Then*

$$E^{\mu_{\Lambda,n}}[\{\sqrt{f(\tilde{T}_{xy}\eta)}-\sqrt{f(\eta)}\}^2] \le CL\tilde{D}^{\gamma_{xy}}(\sqrt{f})$$

Lemma 5.1 can be proved using exactly the same procedure of [KOV]. The idea can be summarized as follows. Since our goal is to exchange configurations at two sites x, y along a path, it is most nature to accomplish this by first exchanging configurations at $x = \gamma_0$ and γ_1 and then exchanging configurations at γ_1 and γ_2 etc, until exchanging configurations at γ_{s-1} and $\gamma_s = y$. Here we have assumed $\gamma_s = y$. Now the configuration at γ_{s-1} is the old configuration at y and we can move this configuration to x using similar procedure. The number of exchanges is $2(s+1) - 3$. Lemma 5.1 follows from the Schwarz inequality.

Return to the proof of Lemma 4.3. Our goal is to move a particle from x to y. To achieve this, let us first exchange configurations at x and γ_{s-1}. This can be done using the Dirichlet form in Lemma 5.1. Now the new configuration at γ_{s-1} is the old configuration at x. Clearly we can move a particle from γ_{s-1} to y, for otherwise Lemma 4.3 is trivial. After moving a particle to y, we can then move the new configuration at γ_{s-1} back to x. Using the Schwarz inequality again, we proved Lemma 4.3.

6 Large Deviation and Two-Block Estimate

We now prove Lemma 4.4. Our basic tools are the local limit theorem and a precise "two-block estimate". The two-block estimate is a standard method in hydrodynamical limit [GPV , KOV]. Our bound gives also an error estimate.

Throughout this section, Λ_L is a cube of width L with volume $L^d = N$, $\mu^c_{\Lambda_L,n} = \mu^c_{L,n}$ is the canonical measure defined in Sect 2 and $y = n/N$ is the density. Define the measure $\mu^{gc}_{L,\lambda}$ be the product measure with marginal density proportional to $\exp[\lambda\eta]\bar\mu(\eta)$, where $\bar\mu$ is the marginal of μ^{gc}_L defined in sect. 2. We shall sometimes use $\langle\cdot\rangle_n$ to denote the expectation with respect to $\mu^c_{L,n}$. Suppose f is a probability density relative to $\mu^c_{L,n}$. Let $s(f)$ be the (specific) entropy of f with respect to $\mu^c_{L,n}$ defined by $s(f) = N^{-1}E^{\mu^c_{L,n}}[f\log f]$. The cube Λ_L is assumed to be a square cube. Certainly, one can generalize it to rectangle without difficulty. We shall also assume that there is an $u(L)$ such that for any probability density f

$$\ell^{-d}E^{\mu^c_{\ell,n}}[f\log f] \le u(L)\ell^2 Av_{b\in\Lambda_\ell} D_b(\sqrt{f}) \tag{6.1}$$

for all $\ell \le L$. Recall the entropy bound:

$$E^{\mu^c_{L,n}}[fX] \le \beta^{-1}L^{-d}\log\int \exp[\beta L^d X]d\mu^c_{L,n} + \beta^{-1}s(f). \tag{6.2}$$

for all $\beta > 0$. This bound is just a consequence of the Jensen inequality but is very useful in large deviation.

The following theorem is version of local limit theorem [P]. We shall only outline its proof. Because of the particle-hole duality, without loss of generality we assume that $y \le R/2$. *This convention will be assumed throughout this section.* Recall $\bar\eta = Av_{x\in\Lambda}\eta_x$.

Theorem 6.1 *Let*

$$Z^c_{L,n} = E^{\mu^{gc}_{L,\lambda}}[\bar\eta = n/N] \tag{6.3}$$

where the chemical chemical potential λ is chosen such that

$$E^{\mu^{gc}_{L,\lambda}}[\bar\eta] = n/N = y.$$

Suppose $L^{-d/2-\alpha} \le y \le R - L^{-d/2-\alpha}$ for some small positive constant α. Then

$$\log Z^c_{L,n} = C_N - \log\sigma + \frac{\gamma_N}{8} N^{-1}(\frac{m^4}{\sigma^4} - 3) + O(N^{-2+3\alpha} y^{-2}) \tag{6.4}$$

where σ^2 is the variance of $\eta_x - y$, m^4 is the 4-th moment of $\eta_x - y$ and C_N and γ_N are universal constants depending only on N such that $\lim_{N\to\infty}\gamma_N = 1$

Proof. For notational simplicity, we shall only prove Theorem 6.1 for the case of Bernoulli random variables (i.e., $R = 1$) and assume $y \le 1/2$. The general case can be proved in the same way. We also assume that N is odd.

Step 1. Fourier Transform. Recall the elementary identity of Fourier series,

$$\frac{1}{N}\sum_{k=0}^{N-1} e^{i2\pi kx/N} = 1 \text{ if } x = 0 \text{ , and } = 0 \text{ otherwise}$$

where

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

Hence

$$Z^c_{L,n} = E^{\mu^{gc}_{L,\lambda}}[\frac{1}{N}\sum_{k=0}^{N-1} e^{i2\pi k(\bar\eta - (n/N))}].$$

Step 2. Cutoff for $k > y^{-1/2}N^{(1+\alpha)/2}$. By periodicity,

$$\frac{1}{N}\sum_{k=0}^{N-1} e^{i2\pi k(\bar\eta-(n/N))} = \frac{1}{N}\sum_{|k|<N/2} e^{i2\pi k(\bar\eta-(n/N))}.$$

Let

$$\Omega^{(\alpha)} = \frac{1}{N}\sum_{y^{-1/2}N^{(1+\alpha)/2}<|k|<N/2} e^{i2\pi k(\bar\eta-(n/N))}$$

Recall that $n/N = y = E^{\mu^{gc}_{L,\lambda}}[\eta_x]$ denotes the probability $\eta_x = 1$. Since $\mu^{gc}_{L,\lambda}$ is a product measure,

$$\begin{aligned} E^{\mu^{gc}_{L,\lambda}}\left\{\prod_x \exp[i2\pi k(\eta_x - y)/N]\right\} &= \prod_x E^{\mu^{gc}_{L,\lambda}}\{\exp[i2\pi k\eta_x/N]\}\exp[-i2\pi ky/N] \\ &= \prod_x \left\{ye^{i2\pi k/N} + (1-y)\right\}\exp[-i2\pi ky/N], \end{aligned}$$

Recall the elementary bound that for any $0 \leq a \leq 1$ and $-\pi \leq \theta \leq \pi$ one has

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

Let $\tilde{k} = 2\pi k$. Hence,

$$\begin{aligned} |E^{\mu^{gc}_{L,\lambda}}[\Omega^{(\alpha)}]| &\leq N^{-1} \sum_{y^{-1/2}N^{(1+\alpha)/2}<|k|<N/2} \exp(-\sum_{x\in\Lambda_L} \frac{y(1-y)\tilde{k}^2}{N^2}) \\ &\leq \sum_{y^{-1/2}N^{(1+\alpha)/2}<|k|<N/2} \exp(-\frac{y(1-y)\tilde{k}^2}{N}) \end{aligned}$$

Therefore, from the range of k one has

$$|E^{\mu^{gc}_{L,\lambda}}[\Omega^{(\alpha)}]| \leq \exp(-CN^{\alpha/2}).$$

Hence the contribution in this region is negligible.

Step 3. Perturbation Expansion. From the cutoff and the definition of $Z^c_{L,n}$, our goal is to estimate

$$\log Z^c_{L,n} = \log\left\{\frac{1}{N}\sum_{|k|\leq y^{-1/2}N^{(1+\alpha)/2}} E^{\mu^{gc}_{L,\lambda}}\left[\exp(i\frac{\tilde{k}}{N}\sum_x(\eta_x - y))\right]\right\} + \exp(-CN^{\alpha/2}).$$

Since $\mu^{gc}_{L,\lambda}$ is a product measure, the expectation can be computed as

$$E^{\mu^{gc}_{L,\lambda}}\left[\exp\{i\frac{\tilde{k}}{N}\sum_x(\eta_x-y)\}\right] = \prod_x E^{\mu^{gc}_{L,\lambda}}\left[\exp\{i\frac{\tilde{k}}{N}(\eta_x-y)\}\right]$$

From the assumption $y \geq L^{-d/2-\alpha}$ and $|k| \leq y^{-1/2}N^{(1+\alpha)/2}$ we have $\tilde{k}/N \ll 1$. Hence we can now expand the exponential to have

$$\begin{aligned} & E^{\mu^{gc}_{L,\lambda}}\left[\exp\{i\frac{\tilde{k}}{N}(\eta_x-y)\}\right] \\ = \;& 1-\frac{1}{2}(\tilde{k}/N)^2\sigma^2 - i\frac{1}{6}(\tilde{k}/N)^3m^3 + \frac{1}{24}(\tilde{k}/N)^4m^4 + O((\tilde{k}/N)^5)y \\ = \;& \exp\left[-\frac{\sigma^2\tilde{k}^2}{2N^2}\right]\left[1 - i\frac{1}{6}(\tilde{k}/N)^3m^3 + \frac{1}{24}(\tilde{k}/N)^4(m^4-3\sigma^4) + O((\tilde{k}/N)^5)y\right] \end{aligned}$$

where σ is the variance of η_x and m^j is the j-th moment of $\eta_x - y$. Note that $m^j \leq$ const. y and we have used this fact in the fifth order term. Using this bound and summing over x and $|k| \leq y^{-1/2}N^{(1+\alpha)/2}$, one has

$$\begin{aligned} \log Z^c_{L,n} = \log\Bigg\{ & \frac{1}{N}\sum_{|k|\leq y^{-1/2}N^{(1+\alpha)/2}} \exp\left[-\frac{\sigma^2\tilde{k}^2}{2N}\right] \\ & \times\left[1-i\frac{1}{6}(\tilde{k}/N)^3m^3 + \frac{1}{24}(\tilde{k}/N)^4(m^4-3\sigma^4) + O((\tilde{k}/N)^5)\,y\right]^N\Bigg\} \end{aligned}$$

Note that the cubic term and the fifth order term in $\tilde{k}$ vanishes due to the symmetry in the summation of $\tilde{k}$. Hence the error term is the sixth order term which is of order $NO((\tilde{k}/N)^5)\, y \le O(N^{-2+3\alpha}y^{-2})$. Summing over k, one has

$$\log Z^c_{L,n} = C_N - \log\sigma + \frac{\gamma_N}{8}N^{-1}(\frac{m^4}{\sigma^4} - 3) + O(N^{-2+3\alpha}y^{-2})$$

where C_N is a universal constant and

$$\gamma_N = \frac{\sigma^4}{3\sqrt{2\pi}\sigma}\sum_{|k|\le y^{-1/2}N^{(1+\alpha)/2}} \exp[-(\sigma^2\tilde{k}^2/2N)](\tilde{k}^4/N^2) \to 1, \text{ as } N\to\infty$$

This concludes our proof.

Recall $\bar{\mu}$ is the marginal of μ^{gc}_L. The following Lemma is an corollary of Theorem 6.1. Recall $y = n/N$ is the density.

Lemma 6.2 *Suppose F is a bounded function of η and $L^{-d/2-\alpha} \le y \le R - L^{-d/2-\alpha}$ for some small positive constant α.Then there is a β_0 such that for $|\beta| < \beta_0$ the following bound holds:*

$$\begin{aligned} & \beta^{-1}N^{-1}\log\int \exp[\beta\sum_x F(\eta_x)]\, d\mu^c_{L,n} \\ = \;& \hat{F}(y) + \beta\, G_2(F,y) + G_1(F,y)N^{-1} \\ & +O(\beta^2) + O(\beta N^{-1}) + O(N^{-2}) + O(\beta^{-1}N^{-3+3\alpha}y^{-2})\ , \end{aligned} \tag{6.5}$$

where

$$\begin{aligned} G_1(F,y) &= \frac{1}{2}[\langle\eta;\eta;\eta\rangle_{\bar{\mu}_\lambda}\langle\eta;F\rangle_{\bar{\mu}_\lambda}\langle\eta;\eta\rangle^{-2}_{\bar{\mu}_\lambda} - \langle\eta;\eta;F\rangle_{\bar{\mu}_\lambda}\langle\eta;\eta\rangle^{-1}_{\bar{\mu}_\lambda}] \\ G_2(F,y) &= \frac{1}{2}[\langle F;F\rangle_{\bar{\mu}_\lambda} - \langle\eta;F\rangle^2_{\bar{\mu}_\lambda}\langle\eta;\eta\rangle^{-1}_{\bar{\mu}_\lambda}] \end{aligned} \tag{6.6}$$

Here $\hat{F}$ and the expectation $\langle\ \ \rangle_{\bar{\mu}_\lambda}$ are defined by

$$\hat{F}(y) := \langle F\rangle_{\bar{\mu}_\lambda} := z(\lambda)^{-1}\int F(\eta)\exp(\lambda\eta)\, d\bar{\mu}(\eta)\ , \quad \lambda = h'(y)\ , \tag{6.7}$$

where $h(y)$ is the free energy defined by

$$h(y) = \sup_\theta[\theta y - p(\theta)],\ p(\theta) = \log z(\theta) = \log\int\exp[\theta\eta]d\bar{\mu}(\eta)\,. \tag{6.8}$$

As usual, $\langle\ ;\ \rangle_{\bar{\mu}_\lambda}$ denotes the truncated (or corrected) correlation function.

Remark. The expectations in (6.6) and (6.7) are with respect to the marginal probability and η denotes the configuration of one site.

Proof. From change of variables $\eta_x \to \eta_x - y$, we can assume that $E^{\bar{\mu}}[\eta_x] = y = 0$ and thus $\lambda = 0$ in (6.7). Also, by subtracting a constant we assume $E^{\bar{\mu}}[F(\eta)] = 0$. Choose γ such that

$$\int \eta\exp(\gamma\eta + \beta F(\eta)\,)\, d\bar{\mu}(\eta) = 0\ . \tag{6.9}$$

Clearly, γ can be solved as

$$\gamma = -\beta\langle\eta F(\eta)\rangle\,\langle\eta^2\rangle^{-1} + \mathrm{O}(\beta^2)\ , \tag{6.10}$$

where $\langle\quad\rangle$ denotes expectation with respect to the density $\bar{\mu}$. Let ξ be the random variable with density relative to $\bar{\mu}$ given by

$$\exp(\gamma\xi + \beta F(\xi)\,)\,z(\beta,\gamma)^{-1} \quad z(\beta,\gamma) = \int \exp(\gamma\eta + \beta F(\eta)\,)d\mu(\eta)\ , \tag{6.11}$$

where $z(\beta,\gamma)$ is the normalization and γ is given by (6.10). By (6.10) and (6.11) we have

$$\begin{aligned} z(\beta,\gamma) &= 1 + \beta^2 G_2(F,0) + O(\beta^3) \\ \sigma_\xi^2 &= z(\beta,\gamma)\sigma^2\{1 - 2\beta G_1(F,0) + O(\beta^2)\} \\ m_\xi^4 &= z(\beta,\gamma)m^4\{1 + O(\beta)\} \end{aligned} \tag{6.12}$$

Let $\{\xi_j\}, j = 1,\cdots,N$ be the i.i.d. with density given by (6.11) and let ν_ξ denote their underlying probability measure. Let

$$Z^c(\beta,\gamma)_{L,n} = E^{\nu_\xi}[\bar{\xi} = n] \tag{6.13}$$

be the partition function. Then we can express the expectation in (6.5) as

$$\beta^{-1}N^{-1}\log\int \exp[\beta\sum_x F(\eta_x)]d\mu^c_{L,n} = \beta^{-1}\log z(\beta,\gamma) + \beta^{-1}N^{-1}[\log Z^c(\beta,\gamma)_{L,n} - \log Z^c_{L,n}]$$

From (6.12), we have

$$\beta^{-1}\log z(\beta,\gamma) = \beta G_2(F,0) + O(\beta^2)$$

From Theorem 6.1,

$$\begin{aligned} & \beta^{-1}N^{-1}\left[\log Z^c(\beta,\gamma)_{L,n} - \log Z^c_{L,n}\right] \\ = \ & \beta^{-1}N^{-1}\left[-\log\sigma_\xi + \log\sigma\right] + \frac{\gamma_N}{8}\beta^{-1}N^{-2}\left\{\frac{m_\xi^4}{\sigma_\xi^4} - \frac{m^4}{\sigma^4}\right\} + O(\beta^{-1}N^{-3+3\alpha}y^{-2}) \end{aligned}$$

From (6.12)

$$\begin{aligned} \beta^{-1}N^{-1}[-\log\sigma_\xi + \log\sigma] &= -\frac{1}{2}\beta^{-1}N^{-1}\log z(\beta,\gamma) - \beta^{-1}N^{-1}\{1 - 2\beta G_1(F,0) + O(\beta^2)\} \\ &= N^{-1}G_1 + O(\beta N^{-1}) \end{aligned}$$

Again, from (6.12)

$$\frac{\gamma_N}{8}\beta^{-1}N^{-2}\left\{\frac{m_\xi^4}{\sigma_\xi^4} - \frac{m^4}{\sigma^4}\right\} = O(N^{-2})$$

We have thus proved lemma 6.2.

Corollary 6.3 *Define* $\tilde{F}_N$ *by*

$$\tilde{F}_N(y) = \int N^{-1} \sum_{x\in\Lambda_L} F(\eta_x)\, d\mu^c_{L,n} \ . \tag{6.14}$$

Then

$$\tilde{F}_N(y) = \hat{F}(y) + G_1(F,y)N^{-1} + O(N^{-2+4\alpha}y^{-2}) + O(N^{-1-\alpha}) \ . \tag{6.15}$$

where $\tilde{F}(y)$ *is defined in (6.7, 6.8).*

Proof By convexity we have for all $\beta > 0$

$$-\beta^{-1}N^{-1}\log\int \exp\,[-\beta\sum_x F(\eta_x)]\, d\mu^c_{L,n} \le \tilde{F}_N(y)$$
$$\le \beta^{-1}N^{-1}\log\int \exp\,[\beta\sum_x F(\eta_x)]\, d\mu^c_{L,n}$$

Choose $\beta = N^{-1-\alpha}$ and use Lemma 6.2, we have Corollary 6.3.

Note that the error terms in (6.15) blow up if $y \le N^{-1+\alpha}$. The following Lemma is useful for the cutoff of low density region.

Lemma 6.4 *Suppose* g *is a function satisfying for some* $\alpha > 0$ *that*

$$\| g - v \|_\infty \le L^{-(d/2)-\alpha}$$

where v *denotes the expectation of* g *with respect to a probability measure* ν*. Then for any* $\beta > 0$*,*

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

and $\delta \to 0$ *as* $L \to \infty$*.*

Proof of Lemma 6.4. *Case 1*: $\beta \ge L^{-(d+\alpha)/2}$. From the assumption on g, we can bound g by $\| g - v \|_\infty \le L^{-(d/2)-\alpha}$. Hence

$$\beta^{-1}L^{-d}\log E^\nu\{\exp[\beta L^d(g-v)]\} \le \| g - v \|_\infty \le L^{-(d/2)-\alpha} \le \delta\beta.$$

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

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

The higher order orders terms are even smaller. This concludes Lemma 6.4.

Lemma 6.5. *For any probability density* f *and any bounded function* F*, one has*

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le Cs(f) \tag{6.16}$$

for some constant C. Recall $y = n/N$ is the density. Then for any constant $\delta > 0$ there is a $y_0 > 0$ independent of L such that for $y \le y_0$

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le \delta s(f) \tag{6.17}$$

provided L is sufficiently large.

Proof. From the entropy bound for any $\beta > 0$,

$$\langle f; Av_x F(\eta_x)\rangle_n \le Q_\beta + \beta^{-1} s(f), \tag{6.18}$$

where $s(f) = s(f/\mu_{L,n}^c)$ and

$$Q_\beta = \beta^{-1} L^{-d} \log \int \exp[\beta \sum_x (F(\eta_x) - q)] d\mu_{L,n}^c$$

and $q = \langle F(\eta_x)\rangle_n$. Suppose $y \le N^{-(1/2)-\alpha}$. Without loss of generality we can assume that $F(\eta_x = 0) = 0$. Hence

$$\| Av_x(F(\eta_x) - q) \|_\infty \le Cy \le CN^{-1/2-\alpha}.$$

Applying Lemma 6.4, we have for any $\beta > 0$, $Q_\beta \le \delta\beta$. Hence

$$\langle f; Av_x F(\eta_x)\rangle_n \le \delta\beta + \beta^{-1} s(f).$$

Optimizing β, we have $\beta = (s(f)/\delta)^{1/2}$ and

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le 2\delta s(f).$$

This proves (6.18) if $y \le N^{-(1/2)-\alpha}$

We now prove (6.16) and (6.17) assuming $y > N^{-(1/2)-\alpha}$. We claim that there is a β_0 such that for any $0 < \beta \le \beta_0$ the follow bound holds:

$$Q_\beta \le C(y)\beta \tag{6.19}$$

for some bounded constant $C(y)$. Furthermore, there is a y_0 such that

$$C(y) \le \delta, \quad \text{if } y \le y_0. \tag{6.20}$$

Assuming these estimates, we continue the proof of Lemma 6.5. Choose $\beta = \min\{\beta_0, \sqrt{s(f)}\}$. By definition of the specific entropy, $s(f)$ is uniformly bounded. Hence from (6.18), (6.19) and (6.20), we have

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le \begin{cases} 2C(y)s(f), & \text{if } \sqrt{s(f)} \le \beta_0 \\ Cs(f) & \text{if } \sqrt{s(f)} \ge \beta_0 \end{cases}$$

This proves (6.16). We now prove (6.17). It suffices to prove that

$$\lim_{y\to 0} s(f/\mu^c_{L,n}) = 0$$

To prove this bound, we have, by definition of entropy,

$$s(f/\mu^c_{L,n}) \le N^{-1} \max_{\eta:|\eta|=n} \log f(\eta) \le -N^{-1} \log \min_{\eta:|\eta|=n} P^{\mu^c_{L,n}}[\eta]$$

Clearly,

$$-N^{-1} \log \min_{\eta:|\eta|=n} P^{\mu^c_{L,n}}[\eta] \le CN^{-1} \log \binom{N}{n} \le -C\left[y \log y + (1-y)\log(1-y)\right]$$

Hence $\lim_{y\to 0} s(f/\mu^c_{L,n}) = 0$. Therefore, $\sqrt{s(f)} \le \beta_0$ and (6.17) holds assuming (6.19) and (6.20).

Finally we have to prove (6.19) and (6.20). Suppose $\beta \ge \epsilon N^{-1} y^{-1}$ for some small constant ϵ. Hence

$$\beta \parallel \sum_x (F(\eta_x) - q) \parallel_\infty \le CN\beta y \le C\epsilon.$$

We can expand the exponential to prove (6.19). Hence we can assume that $\epsilon N^{-1} y^{-1} \le \beta \le \beta_0$ and $y \ge N^{-(1/2)-\alpha}$. Under this assumption, the error terms in (6.5) can be bounded by

$$C\beta\left[\beta + N^{-1} + N^{-2}\beta^{-1} + \{N^{-1}\beta^{-1}y^{-1}\}^2 N^{-1+3\alpha}\right] \le C\beta[\beta + L^{-\alpha}]$$

Hence from (6.5)

$$Q_\beta \le \hat{F}(y) + \beta\, G_2(F,y) + G_1(F,y)N^{-1} - q + O\left(\beta[\beta + L^{-\alpha}]\right)$$

From (6.15)

$$\begin{aligned} \hat{F}(y) + G_1(F,y)N^{-1} - q &\le O(N^{-2+4\alpha}y^{-2}) + O(N^{-1-\alpha}) \\ &\le C\beta\left[(\beta^{-1}N^{1}y^{-1})N^{-1+4\alpha}y^{-1} + N^{-1-\alpha}\beta^{-1}\right] \end{aligned}$$

From assumptions $\epsilon N^{-1}y^{-1} \le \beta \le \beta_0$ and $y \ge N^{-(1/2)-\alpha}$ the last term is bounded by $C\beta L^{-\alpha}$. Hence we have proved

$$Q_\beta \le \beta\, G_2(F,y) + O\left(\beta[\beta + L^{-\alpha}]\right)$$

This proves (6.19). Since $G_2(F,y) \to 0$ as $y \to 0$, we have also proved (6.20). This concludes the proof of Lemma 6.5.

If (6.17) holds without the restriction on small y, Lemma 4.4 follows immediately. Unfortunately, we can not prove such a bound directly and a more indirect approach is needed. Note that we have proved Lemma 4.4 if the density y is small. By particle-hole duality, similar result holds for $y \to R$. Therefore, we can assume y is bounded away from the 0 or R. *For the rest of this section, we assume that* $y_0 \le y \le R - y_0$.

Lemma 6.6 *Divide* Λ *into cubes of width* k *and index these cubes by* σ. *Let* $\bar\eta_\sigma$ *be the average of* η *in the cube* σ *and let* $\tilde F_k(\bar\eta_\sigma)$ *be defined by*

$$\tilde F_k(\bar\eta_\sigma) = E^{\mu_L^{gc}}[Av_{x\in\sigma}F(\eta_x)|\bar\eta_\sigma]. \tag{6.21}$$

Then there is a $C_1(k)$ *such that for any probability density* f

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le 2\langle f; Av_\sigma \tilde F_k(\bar\eta_\sigma)\rangle_n^2 + C_1(k) D_\Lambda(\sqrt{f}) \tag{6.22}$$

Proof. From the Schwarz inequality,

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le 2\langle f; Av_\sigma \tilde F_k(\bar\eta_\sigma)\rangle_n^2 + 2Av_\sigma\langle f; Av_{x\in\sigma}\{F(\eta_x) - \tilde F_k(\bar\eta_\sigma)\}\rangle_n^2$$

From the definition of $\tilde F_k$ and the definition of conditional expectation,

$$\langle f; Av_{x\in\sigma}\{F(\eta_x) - \tilde F_k(\bar\eta_\sigma)\}\rangle_n = Av_\sigma E^{\mu^c_{L,n}}\left\{E^{\mu^c_{L,n}}[f; Av_{x\in\sigma}F(\eta_x)|\bar\eta_\sigma]\right\}. \tag{6.23}$$

For each σ fixed, we can now apply lemma 6.5 to bound the right side of (6.22). Note that on a cube of width k the logarithmic Sobolev inequality holds with some constant $C_1(k)$. Hence we can replace the entropy in the σ-th cube by the Dirichlet form $C_1(k)\sum_{b\in\sigma} D_b(\sqrt{f})$. This proves Lemma 6.6.

Lemma 6.6 replace the local function $F(\eta_x)$ by its local average $\tilde F_k(\bar\eta_\sigma)$. The follow lemma proves Lemma 4.4 when $u(L)L^{2-d}D(\sqrt{f})$ is small.

Lemma 6.7 *Suppose* $y_0 \le y \le R - y_0$. *Then for any* $\epsilon_0 > 0$ *there is a* $\delta > 0$ *such that for any probability density* f *with* $u(L)L^{2-d}D(\sqrt{f}) \le \delta$

$$\langle f; Av_x F(\eta_x)\rangle_n^2 \le [\epsilon_0 u(L) + C(k)]L^{2-d}D(\sqrt{f}).$$

provided that k *is large enough.*

Proof. Let

$$W_k(z) = \tilde F_k(z) - a - b(z-y),$$
$$\Omega_\beta = (\beta k^d)^{-1}(L/k)^{-d}\log\int \exp[(\beta k^d)(L/k)^d Av_\sigma\{W_k(\bar\eta_\sigma)\}d\mu^c_{L,n} \tag{6.24}$$

From the entropy bound, for any $\beta > 0$,

$$\langle f; Av_\sigma \tilde F_k(\bar\eta_\sigma)\rangle_n = \langle f; Av_\sigma W_k(\bar\eta_\sigma)\rangle_n \le \Omega_\beta + s/\beta, \tag{6.25}$$

where we have added $Av_\sigma a + b(\bar\eta_\sigma - y) = a$, which is simply a constant (with respect to $\mu^c_{L,n}$) and can be added at will. The choice of a, b will be made later. As in the proof of Lemma 6.5, our goal is to prove

$$\Omega_\beta \le \epsilon_0 \beta, \quad \text{if } \beta \le \delta^{1/3} \tag{6.26}$$

Assuming this bound, we have from (6.25),

$$\langle f; Av_\sigma \tilde F_k(\bar\eta_\sigma)\rangle_n \le \epsilon_0\beta + \beta^{-1}\epsilon_0 u(L) L^{2-d} D(\sqrt{f}),$$

where we have replaced the entropy by the Dirichlet form using the definition of $u(L)$. Optimizing β, we have $\beta = [u(L)L^{2-d}D(\sqrt{f})/\epsilon_0]^{1/2}$ and

$$\langle f; Av_\sigma \tilde F_k(\bar\eta_\sigma)\rangle_n^2 \le 2\epsilon_0 u(L) L^{2-d} D(\sqrt{f})$$

provided that $[u(L)L^{2-d}D(\sqrt{f})/\epsilon_0]^{1/2} \le \delta^{1/3}$. From the assumption on the Dirichlet form, $[u(L)L^{2-d}D(\sqrt{f})/\epsilon_0]^{1/2} \le (\delta/\epsilon_0)^{1/2}$. Hence for δ small enough, the optimal β satisfies $\beta \le \delta^{1/3}$. Together with Lemma 6.6, this proves Lemma 6.7 assuming (6.26).

We now prove (6.26). By expanding the exponential, (6.26) holds if $\beta \le C\sqrt{\epsilon_0}N^{-1}$. Hence we can assume that $C\sqrt{\epsilon_0}N^{-1} \le \beta \le \delta^{1/3}$.

Note that, with respect to $\mu^c_{L,n}$, $\bar\eta_\sigma$ is a family of independent random variable distributed according to μ^{gc}_L with global constraint $Av_\sigma \bar\eta_\sigma = y$. More precisely, $\bar\eta_\sigma$ is characterized by

$$\langle g(\bar\eta_k)\rangle_{\mu^{gc}_k} = \int g(\bar\eta_k)\, d\mu^{gc}_k .$$

From Lemma 6.2 and the assumption $y_0 \le y \le R - y_0$,

$$\Omega_\beta = \langle W_k\rangle_{\mu^{gc}_{k,\lambda}} + G_1(W_k, y)k^d N^{-1} + \beta k^d\, G_2(W_k, y) + \beta O(\beta + L^{-\alpha})\ , \tag{6.27}$$

where G_1 and G_2 are defined in (6.6) with η replaced by $\bar\eta_k$, i.e.,

$$G_2(W_k, y) = \frac{1}{2}[\langle W_k; W_k\rangle_{\mu^{gc}_{k,\lambda}} - \langle \bar\eta; W_k\rangle^2_{\mu^{gc}_{k,\lambda}} \langle \bar\eta; \bar\eta\rangle^{-1}_{\mu^{gc}_{k,\lambda}}]$$

and similarly for G_1. Here

$$\langle g(\bar\eta_k)\rangle_{\mu^{gc}_{k,\lambda}} = z(\lambda)^{-k^d} \int f(\bar\eta_k) \exp(\lambda k^d \bar\eta_k\,)\, d\mu\ , \quad \lambda = h'(y)\ ,$$

where $h(y)$ is the free energy defined by (6.8) and z is the partition function defined in (6.7). From the restriction on β, for δ small enough we can bound the error term by $\epsilon_0\beta/3$. Choose a such that $\langle W_k\rangle_{\mu^{gc}_{k,\lambda}} + G_1(W_k, y)k^d N^{-1} = 0$. It remains to prove

$$\lim_{k\to\infty} k^d\, G_2(W_k, y) = 0. \tag{6.28}$$

From Corollary 6.3 one has $|\hat{F}-\tilde{F}_k| \leq k^{-d-\alpha}$. Hence we can replace $\tilde{F}_k$ by $\hat{F}$. Let $b=\partial\hat{F}(z)/\partial z\,|_{z=y}$. From the Taylor expansion,

$$\hat{F}(z)-b(z-y)=F(y)+\frac{1}{2}\left(\frac{\partial^2\hat{F}(z)}{\partial z^2}\right)_{z=y}(z-y)^2+\text{ higher order terms}$$

Since $\mu^{gc}_{k,\lambda}$ is a product measure, $\bar{\eta}_k-y\sim O(k^{-d/2})$. Because the linear term vanishes, it is now easy to check

$$\langle\hat{F}(\bar{\eta}_k)-b(\bar{\eta}_k-y);\bar{\eta}_k\rangle^2_{\mu^{gc}_{k,\lambda}}\sim o(k^{-d})\sim\langle\hat{F}(\bar{\eta}_k)-b(\bar{\eta}_k-y);\hat{F}(\bar{\eta}_k)-b(\bar{\eta}_k-y)\rangle_{\mu^{gc}_{k,\lambda}}.$$

This proves (6.26) and concludes Lemma 6.7.

We have proved Lemma 4.4 in case that $u(L)L^2D_\Lambda/L^d$ is small or the density y is small. The next lemma will be used to close this gap. Lemma 4.4 is just a corollary of Lemmas 6.5, 6.7 and 6.8.

Lemma 6.8 *Suppose that $y_0\leq y\leq R-y_0$ and $u(L)L^{2-d}D(\sqrt{f})\geq\delta$ for some probability density f. Then for any $\epsilon_1>0$ and k large enough,*

$$\langle f;Av_\alpha\tilde{F}_k(\bar{\eta}_\alpha)\rangle^2_n\leq[C+\epsilon_1u(L)]L^2Av_{b\in\Lambda}D_b(\sqrt{f})\tag{6.29}$$

Here the constant depends on ϵ_1, δ and F.

We need the following lemma to prove Lemma 6.8. For those familiar with hydrodynamical limit, it can be understood as a statement of two-block estimate.

Lemma 6.9 *Suppose (η,ζ) are configurations on $\Lambda_k\cup\Lambda_k$. Define the Dirichlet form which describes the exchange of particles between η and ζ by*

$$D_{\eta,\zeta}(f)=\sum_x D^{(x)}_{\eta,\zeta}(f)=\sum_x\sum_{\vartheta=\pm}\int[f(\omega^\vartheta_x(\eta,\zeta)-f(\eta,\zeta)]^2d\mu(\eta)d\mu(\zeta)\tag{6.30}$$

Here $\omega^+_x(\eta,\zeta)=(\omega^+_x\eta,\omega^+_x\zeta)$,

$$(\omega^+_x\eta)_z=\begin{cases}\eta_x-1 & \textit{if } z=x,\eta_x>0\textit{ and }\zeta_x<R\\ \eta_z & \textit{otherwise},\end{cases}\tag{6.31}$$

and

$$(\omega^+_x\zeta)_z=\begin{cases}\zeta_x+1 & \textit{if } z=x,\eta_x>0\textit{ and }\zeta_x<R\\ \zeta_z & \textit{otherwise}.\end{cases}$$

Similarly we can define ω^-_x, namely

$$(\omega^-_x\eta)_z=\begin{cases}\eta_x+1 & \textit{if } z=x,\eta_x<R\textit{ and }\zeta_x>0\\ \eta_z & \textit{otherwise},\end{cases}\tag{6.32}$$

and

$$(\omega_x^- \zeta)_z = \begin{cases} \zeta_x - 1 & \text{if } z = x, \eta_x < R \text{ and } \zeta_x > 0 \\ \zeta_z & \text{otherwise .} \end{cases}$$

Then there are constants $C_1(k)$ and C_2 such that for any probability density f

$$\int f(\bar\eta - \bar\zeta)^2 d\mu \le C_1(k) D_{\eta,\zeta}(\sqrt{f}) + C_2/k^d. \tag{6.33}$$

Proof of Lemma 6.9. For any function $g(\eta, \zeta)$ independent of η_x and ζ_x, we have

$$\int g(\eta,\zeta)(\eta_x - \zeta_x) f d\mu \le C D^{(x)}_{\eta,\zeta}(\sqrt{f})$$

where C is a constant depending on $\|g\|_\infty$. To prove this bound, we only have to check it for one site since $g(\eta,\zeta)$ is independent of η_x and ζ_x. But this is obvious because the number of particles per site, R, is finite. Hence

$$\begin{aligned} & \int (\bar\eta - \bar\zeta)(\eta_x - \zeta_x) f d\mu \\ = & \int \left[(\bar\eta - (\eta_x/k^d)) - (\bar\zeta - (\zeta_x/k^d)) \right] (\eta_x - \zeta_x) f d\mu + \int k^{-d}(\eta_x - \zeta_x)(\eta_x - \zeta_x) f d\mu \\ \le & \; C D^{(x)}_{\eta,\zeta}(\sqrt{f}) + C/k^d. \end{aligned}$$

Averaging over $x \in \Lambda_k$,

$$\int (\bar\eta - \bar\zeta)^2 f d\mu \le C Av_x D^{(x)}_{\eta,\zeta}(\sqrt{f}) + C/k^d.$$

This concludes Lemma 6.9.

Proof of Lemma 6.8 For any two cubes σ, τ of width k in Λ one has from Lemma 6.9

$$\int f(\bar\eta_\sigma - \bar\eta_\tau)^2 d\mu^c_{L,n} \le C_1(k) D_{\sigma,\tau}(\sqrt{f}) + C_2(k)$$

Here $D_{\sigma,\tau}$ is defined in (6.30) with the cubes σ, τ take the roles of $\Lambda_k \cup \Lambda_k$. By definition $D_{\sigma,\tau}$ involves exchanges of particles in the cubes σ and τ. From Lemma 4.3, we can bound these exchanges by the usual Dirichlet form with only nearest neighbor exchanges. Average over σ and τ in Λ and then use the Schwarz inequality,

$$Av_\sigma \int f(\bar\eta_\sigma - y)^2 d\mu^c_{L,n} \le C_1(k) L^{2-d} D_\Lambda(\sqrt{f}) + C_2(k)$$

From Corollary 6.3 and the assumption on y,

$$|\tilde F_k(z) - \hat F(z)|^2 \le C k^{-d}$$

if $|z-y| \le y_0/4$. Since $\hat{F}$ is Lipschitz, this implies

$$|\tilde{F}_k(\bar{\eta}_\sigma) - \hat{F}(y)|^2 \le C(\bar{\eta}_\sigma - y)^2 + Ck^{-d}$$

if $|\bar{\eta}_\sigma - y| \le y_0/4$. On the other hand, the last bound holds trivially if $|\bar{\eta}_\sigma - y| \ge y_0/2$ by choosing c depending on y_0. Therefore, we have

$$\begin{aligned} \langle f; Av_\sigma \tilde{F}_k(\bar{\eta}_\sigma)\rangle^2_{\mu^c_{L,n}} &\le Av_\sigma E^f[(\tilde{F}_k(\bar{\eta}_\sigma) - \hat{F}(y))]^2 \le C Av_\sigma E^f[(\bar{\eta}_\sigma - y)^2] + Ck^{-d} \\ &\le C[C_1(k) L^{2-d} D_\Lambda(\sqrt{f}) + k^{-d}] \end{aligned}$$

Choosing k large enough and using the assumption $u(L)L^{2-d}D(\sqrt{f}) \ge \delta$, we have proved lemma 6.8.

Acknowledge. It is a great pleasure to thank S.L. Lu for many discussions regarding the logarithmic Sobolev inequality and spectral gap for the Kawasaki dynamics.

References

[DGS] 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.

[DSC] P. Diaconis and L. Saloff-Coste, Logarithmic Sobolev inequlaity for finite Markov Chains, preprint.

[DSh] P. Diaconis and L. Shahshahani, Time to reach stationarity in the Bernoulli-Laplace diffusion model, SIAM Jour. Math. Anal., **18**, 208-218, 1987.

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

[KOV] C. Kipnis, S. Olla and R. S. R. Varadhan, Hydrodynamics and large deviations for simple exclusion process, *Commun. Pure Appl. Math*,**42**, 115–137, 1989.

[LSV] C. landim, S. Sethuraman and S. R. S. Varadhan Spectral gap for zero range dynamics, to appear in Annals of Prob., (1995)

[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.

[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.

[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.