A LOWER BOUND ON THE GROUND STATE ENERGY OF DILUTE BOSE GAS

JI OON LEE AND JUN YIN

Abstract. Consider an N-Boson system interacting via a two-body repulsive short-range potential V in a three dimensional box Λ of side length L. We take the limit $N, L \to \infty$ while keeping the density $\rho = N/L^3$ fixed and small. We prove a new lower bound for its ground state energy per particle

$$\frac{E(N,\Lambda)}{N} \geq 4\pi a\rho[1 - O(\rho^{1/3}|\log\rho|^3)],$$

as $\rho \to 0$, where a is the scattering length of V.

1. Introduction

The properties of the Bose gas have been studied by many authors [5, 2, 7, 8], and since the first experimental observation of Bose-Einstein Condensation in 1995 [1], interest in low temperature Bose gases are renewed [12, 13, 10, 11, 9, 16, 17]. One of the most well-known properties of Bose gas is its ground state energy in the dilute limit. In this low density limit, the leading term of the ground state energy per particle is $4\pi a\rho$, where a is the scattering length of two-body interaction potential and ρ is the density. The upper bound for it was first rigorously proved by Dyson [2], and the lower bound was obtained by Lieb and Yngvason [12]. It is also proved that the leading term is the same in some cases where the interaction potential is not purely non-negative [4, 20].

Lee and Yang [6, 7] first predicted the second order correction to this leading term, which is given by

$$4\pi a\rho\left[1 + \frac{128}{15\sqrt{\pi}}(\rho a^3)^{\frac{1}{2}} + \cdots\right]. \tag{1.1}$$

This calculation used pseudopotential method and binary collision expansion method [5]. Another derivation was later given by Lieb [8] using a self-consistent closure assumption for the hierarchy of correlation functions.

The upper bound of this second order correction is recently obtained in Yau and Yin's work [19]. The matching lower bound, however, has not been proved yet, and in fact, the best result so far is $4\pi a\rho(1 - C\rho^{1/17})$, which was first derived by Lieb and Yngvason [12]. (See also [14] for details.)

Lieb and Yngvason [12] used the "cell method" to find the lower bound for the ground state energy. In this method, they first converted the interaction potential into a soft potential with the expense of kinetic energy. Unlike Dyson [2], however, they did not sacrifice all the kinetic energy but kept a small portion of kinetic energy to apply Temple's inequality. Then, after dividing a large box into smaller cubic cells, perturbation theory was applied in each small cubic cell. In this way, the correct leading term could be obtained.

We improve this approach with several new ideas including a new method called 'box doubling method.' Some notable advantages of our methods are as follows:

- We use a combination of perturbation theory and Temple inequality. Also, we are able to preserve essentially the full kinetic energy after the replacement of the the singular potential by a smooth one.
- Each small cell can be enlarged so that it contains enough number of particles in average.

In this paper we show that the error term becomes $-C\rho^{1/3}|\log\rho|^3$ with the new method. Our main result is stated in Theorem 2.2.

This paper is organized as follows: In Section 2, we state our main theorem, introduce key lemmas, and prove the theorem. In section 3, we show new type of 'cell method', which is equivalent to proving some of the key lemmas. In section 4, we introduce 'box doubling method', which enlarges each cell, hence makes it contain more particles in average. Some technical estimates can be found in section 5.

2. Main result and Key lemmas

2.1. **Definition of the system and main result.** Let V be a non-negative, smooth, spherically symmetric, and compactly supported potential, satisfying $V(x) = 0$ if $|x| > R_0$ for $x \in \mathbb{R}^3$. The zero energy scattering solution φ satisfies

$$\left(-\Delta + \frac{1}{2}V(x)\right)\varphi(x) = 0. \tag{2.1}$$

We normalize φ so that $\lim_{|x|\to\infty}\varphi(x) = 1$. The scattering length a of V is defined by

$$a := \lim_{|x|\to\infty} |x|\,(1-\varphi(x)) \tag{2.2}$$

as in [12].

Hamiltonian of the Bose gas system of N particles in a torus Λ of side length L is given by

$$H_N := -\sum_{j=1}^{N}\Delta_j + \sum_{i<j}^{N} V(x_i - x_j). \tag{2.3}$$

We will consider the thermodynamic limit when N and L approach infinity with the density $\rho = N/L^3$ fixed. Later, we will also let $\rho \to 0$. We will use the notation $\alpha \ll \beta$ if $\alpha/\beta = O(\rho^c)$ or $\alpha/\beta = O(|\log\rho|^{-c})$ for some $c > 0$, and $\alpha \sim \beta$ if $\alpha/\beta = O(1)$. For simplicity, we will also use the notation

$$\mathbf{x}_n = (x_1, x_2, \cdots, x_n), \quad \mathbf{x} = \mathbf{x}_N. \tag{2.4}$$

Throughout the paper, C denotes a constant independent of ρ.

Definition 2.1 (The ground state energy)**.** For given Hamiltonian

$$H_n := -\sum_{j=1}^{n}\Delta_j + \sum_{i<j}^{n} V(x_i - x_j) \tag{2.5}$$

in the three dimensional torus Λ, its ground state energy, $E(n, \Lambda)$ is defined to be

$$E(n,\Lambda) := \inf_{\|\psi\|_2=1}\Big[\sum_{j=1}^{n}\int_{\Lambda^N}|\nabla_j\psi(\mathbf{x})|^2 d\mathbf{x} + \sum_{i<j}^{n}\int_{\Lambda^N}V(x_i-x_j)|\psi(\mathbf{x})|^2 d\mathbf{x}\Big]. \tag{2.6}$$

The main result of this paper is the following theorem.

Theorem 2.2. *Let V be non-negative, smooth, spherically symmetric, and compactly supported potential, satisfying $V(x) = 0$ if $|x| > R_0$, whose scattering length is a. Then, there exists a constant $C_0 > 0$ such that*

$$\lim_{\substack{N,L\to\infty\\ N/L^3=\rho}} \frac{E(N,\Lambda)}{N} \geq 4\pi a\rho(1 - C_0\rho^{\frac{1}{3}}|\log\rho|^3) \tag{2.7}$$

as $\rho \to 0$.

2.2. **Notations.** Let $e_0(\kappa)$ and $(1-\tau(\kappa,x))$ be the lowest Neumann eigenvalue and eigenfunction of $(-\Delta+\frac{1}{2}V)$ on the ball of radius $\kappa \gg R_0$, i.e.,

$$(-\Delta+\frac{1}{2}V(x))(1-\tau(\kappa,x)) = e_0(\kappa)(1-\tau(\kappa,x)) \tag{2.8}$$

with the boundary conditions

$$\tau(\kappa,x)=0, \quad \partial\tau(\kappa,x)=0, \quad \text{if } |x|=\kappa. \tag{2.9}$$

Note that τ is spherically symmetric, since V is spherically symmetric. Some properties of τ and e_0 are collected in Section 5, most notably,

$$\frac{3a}{\kappa^3} \le e_0 \le \frac{3a}{\kappa^3}(1+\frac{C}{\kappa}). \tag{2.10}$$

Using τ, we define a function W_j that will be used for our proof, which will approximate the ground state of $-\Delta_j+\sum_{i:i\neq j} V(x_i-x_j)$ when $\mathbf{x}$ is fixed except x_j. Let t_{ij} be the half of the distance of x_i to its nearest particle other than x_j, i.e.,

$$t_{ij} = \frac{1}{2}\min_{k:k\neq i,j} |x_i-x_k|. \tag{2.11}$$

We introduce a particle triple cutoff function $F_{ij}(\mathbf{x})$ at a length scale ℓ_0 defined by

$$F_{ij}(\mathbf{x}) = \begin{cases} 1 & \text{if } t_{ij} > \ell_0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.12}$$

We also introduce a particle triple cutoff function with a length scale $\ell_{-1} \ll \ell_0$, G_{ij}, which is defined to be

$$G_{ij}(\mathbf{x}) = \begin{cases} 1 & \text{if } t_{ij} > \ell_{-1} \\ 0 & \text{otherwise} \end{cases}. \tag{2.13}$$

Here, ℓ_0 and ℓ_{-1} are fixed and satisfy $\ell_0 \gg \ell_{-1} \gg R_0$. Note that $t_{ij} \neq t_{ji}$, and in fact, t_{ij} does not depend on x_j. The same holds for F_{ij} and G_{ij}.

Let us extend the definition of τ so that $\tau(\kappa,x)=0$ if $|x|>\kappa$. We define W_j by

$$W_j(\mathbf{x}) = 1-\sum_{i:i\neq j}[F_{ij}(\mathbf{x})\tau(\ell_0,x_i-x_j)+(1-F_{ij}(\mathbf{x}))G_{ij}(\mathbf{x})\tau(t_{ij},x_i-x_j)]. \tag{2.14}$$

Our interpretation of this wave function W_j is as follows: When x_i is the nearest particle of x_j, W_j becomes $1-\tau(\ell_0,x_i-x_j)$ or $1-\tau(t_{ij},x_i-x_j)$, depending on t_{ij}. Physically, it corresponds to a situation in which the particle x_j is moving and all the other particles are fixed. Any particle x_i other than x_j has its own 'effect' on x_j, and the radius of this effect is given by ℓ_0 or t_{ij}, provided that $t_{ij} \ge \ell_{-1}$. x_j is affected only by its nearest particle.

The definition of W_j shows how we can handle the case when three or more particles are close to each other; when a particle x_k is close to x_i and x_j, we do not ignore interaction between x_i and x_j but assume that the range of interaction shrinks accordingly. Here, the range of interaction is t_{ij}, the size of support of τ.

We will use length scales from now on such that

$$\ell_{-1} \sim \rho^{-\frac{2}{9}}|\log\rho|^{\frac{2}{3}}, \quad \ell_0 \sim \rho^{-\frac{1}{3}}|\log\rho|^{-\frac{1}{3}}, \quad \ell_1 \sim \rho^{-\frac{1}{3}}|\log\rho|^{\frac{1}{3}+\eta}, \quad \ell_2 \sim \rho^{-\frac{4}{9}}|\log\rho|^{-\frac{2}{3}}, \quad \epsilon = \rho^{\frac{1}{3}}|\log\rho|^3. \tag{2.15}$$

Here, η is a small positive number such that $0<\eta<1/15$. We choose the parameters properly to satisfy

$$L = h'\ell_2, \quad \ell_2 = 2^h\ell_1, \tag{2.16}$$

where h, h' are integers.

We will prove in Section 5 that there exists a constant $c_0 < 1$ such that

$$W_j(\mathbf{x}) \geq 1 - c_0 > 0. \tag{2.17}$$

Other properties of W_j will be proved in Section 5 as well.

2.3. **Basic strategy of the proof.** To prove the main theorem, we use a strategy that consists of four steps:

(1) For a given Bosonic wave function $\Psi \in L^2(\Lambda^N)$, we let $\Phi_j := W_j^{-1}\Psi$ for $j = 1, 2, \cdots, N$. Φ_j is well defined, since $W_j > 0$. Using W_j, we convert the hard potential V into a soft potential q with the expense of a portion of local kinetic energy as we can see from the following lemma:

Lemma 2.3. *Let*

$$q(\kappa, x) := e_0(\kappa) 1(|x| \leq \kappa) \tag{2.18}$$

and

$$q_{ij} := F_{ij} q(\ell_0, x_i - x_j) + (1 - F_{ij}) G_{ij} q(t_{ij}, x_i - x_j). \tag{2.19}$$

Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$. *Then, for any* $\Psi(\boldsymbol{x}) \in L^2(\Lambda^N)$,

$$\langle \Psi, H_N \Psi \rangle \geq \sum_{j=1}^N \int_{\Lambda^N} |W_j|^2 |\nabla_j \Phi_j(\boldsymbol{x})|^2 d\boldsymbol{x} + \sum_{i \neq j}^N \int_{\Lambda^N} q_{ij}(\boldsymbol{x}) |\Psi(\boldsymbol{x})|^2 d\boldsymbol{x}, \tag{2.20}$$

where $\Phi_j(\boldsymbol{x}) := W_j^{-1} \Psi(\boldsymbol{x})$ *for* $j = 1, 2, \cdots, N$.

Note that, though q_{ij} is the soft potential we want to use throughout the paper, it is not symmetric, i.e., $q_{ij} \neq q_{ji}$. The kinetic energy term here contains $|W_j|^2 |\nabla_j \Phi_j(\mathbf{x})|^2$, which is associated with an operator $T_j = -W_j^{-1} \nabla_j W_j^2 \nabla_j W_j^{-1}$. To use this operator T_j, we need to impose boundary condition. Noting that Neumann boundary condition gives the lowest energy for a Laplacian operator $-\Delta_j$, we introduce the following condition on T_j that corresponds to the Neumann boundary conditions on $-\Delta_j$:

Definition 2.4. Let $\Lambda_X \subset \Lambda$ be a box.

(a) Assume that $x_1, \cdots, \widehat{x_j}, \cdots, x_N$ are fixed. We say that $\psi(x_j)$ satisfies 'W_j-Neumann boundary condition' in Λ_X when $\psi(x_j) \in L^2(\Lambda_X)$ and $W_j^{-1}\psi(x_j)$ satisfies Neumann boundary condition in Λ_X.

(b) Let $j = 1, 2, \cdots, n$. Assume that $x_{n+1}, \cdots, x_N$ are fixed. We say that $\psi(\mathbf{x}_n)$ satisfies 'W_j-Neumann boundary condition' in Λ_X when $\psi(\mathbf{x}_n) \in L^2(\Lambda_X^n)$ and, for any fixed $x_1, \cdots, \widehat{x_j}, \cdots, x_n$,

$$\psi_j(x_j) := \psi(\mathbf{x}_n) \tag{2.21}$$

satisfies W_j-Neumann boundary condition in Λ_X as defined above.

When $x_{n+1}, x_{n+2}, \cdots, x_N$ are fixed outside Λ_X, for a given $\psi(\mathbf{x}_n) \in L^2(\Lambda_X^n)$, we can find $\widetilde{\psi}(\mathbf{x}_n) \in L^2(\Lambda_X^n)$ such that $\widetilde{\psi}$ satisfies W_j-Neumann boundary condition in Λ_X for $j = 1, 2, \cdots, n$, and the difference between the kinetic energies of ψ and $\widetilde{\psi}$ is negligible.

(2) To use perturbation theory, we divide the torus Λ into small cubic cells. We need to replace the soft potential q_{ij} by another soft potential $\widetilde{q_{ij}}$ so that particles in different boxes do not interact via this soft potential. If we just ignore interactions beyond the boundaries of cubic cells, however, it will lower the energy significantly. To resolve this problem, we shift the origin of division, u, continuously and take an average of energy with respect to u. This technique leads us to the following lemma:

Lemma 2.5. *Consider a three dimensional grid with the origin $u \in \Lambda$. Assume that this grid divides Λ into small cubic cells of side length $\ell_1 \sim \rho^{-\frac{1}{3}}|\log\rho|^{\frac{1}{3}+\eta}$, $0 < \eta < 1/15$ such that*

$$\Lambda = \bigcup_{\lambda} \Lambda_{u\lambda}, \quad \Lambda_{u\lambda_i} \cap \Lambda_{u\lambda_j} = \emptyset \text{ if } \lambda_i \neq \lambda_j, \tag{2.22}$$

where $\Lambda_{u\lambda}$ is a cubic cell of side length ℓ_1 and λ is an index for those cubic cells.

For given u, let $\partial\Lambda_{u\lambda}$ be the boundary of $\Lambda_{u\lambda}$ and $d(x_i, \partial\Lambda_{u\lambda})$ the distance between x_i and $\partial\Lambda_{u\lambda}$ in Λ. Define

$$\widetilde{q_{ij}} := \begin{cases} q_{ij} & \text{if } d(x_i, \partial\Lambda_{u\lambda}) > 2\ell_0 \text{ for any } \lambda \\ 0 & \text{otherwise} \end{cases} \tag{2.23}$$

Then, for any $\Psi(\boldsymbol{x}) = \Psi(x_1, x_2, \cdots, x_N) \in L^2(\Lambda^N)$,

$$\sum_{i\neq j}^{N} \int_{\Lambda^N} q_{ij}(\boldsymbol{x})|\Psi(\boldsymbol{x})|^2 d\boldsymbol{x} \geq \inf_u (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i\neq j}^{N} \int_{\Lambda^N} \widetilde{q_{ij}}(\boldsymbol{x})|\Psi(\boldsymbol{x})|^2 d\boldsymbol{x}. \tag{2.24}$$

(3) In each small cubic cell of side length ℓ_1, we use perturbation theory with half of kinetic energy to estimate the energy in the small cell in terms of the number of particles in it. When the small cell contains too many particles, however, perturbation loses its validity. In this case, a very small portion of energy from the main Hamiltonian H_N will contribute the energy we need and it is sufficient in order to estimate the lower bound for the ground state energy. More precisely, we can prove the following lemma:

Lemma 2.6. *Let $x_1, x_2, \cdots, x_N \in \Lambda$. Let $B \subset \Lambda$ be a box of side length ℓ_1. Define $E(n, B)$ as the ground state energy of n particle system in B with interaction V, i.e.,*

$$E(n, B) := \inf \text{ spec } \{-\sum_{j=1}^{n} \Delta_j + \sum_{i<j}^{n} V(x_i - x_j)\}. \tag{2.25}$$

Let

$$f(n) := \begin{cases} n(n-1) & \text{if } n \leq 2\rho\ell_1^3 \\ (4\rho\ell_1^3 - 1)n - (2\rho\ell_1^3)^2 & \text{if } n > 2\rho\ell_1^3 \end{cases}. \tag{2.26}$$

Fix $x_{n+1}, x_{n+2}, \cdots, x_N$ outside B and let $\boldsymbol{x}_n = (x_1, x_2, \cdots, x_n)$. Then, for any $\psi(\boldsymbol{x}_n) \in L^2(B^n)$ with $\phi_j := W_j^{-1}\psi$ for $j = 1, 2, \cdots, n$, we have

$$\begin{aligned} &\epsilon E(n, B)\|\psi\|_2^2 + \frac{1}{2}\sum_{j=1}^{n} \int_{B^n} |W_j|^2 |\nabla_j \phi_j(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i\neq j}^{n} \int_{B^n} \widetilde{q_{ij}} |\psi_j(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n \\ \geq \quad & f(n)(1 - C\epsilon)\frac{4\pi a}{\ell_1^3}\|\psi\|_2^2. \end{aligned} \tag{2.27}$$

(4) In the cell method in step (3), the number of pairs in a small cubic cell is not $n^2/2$ but $n(n-1)/2$. This gives an additional factor $(1 - \rho\ell_1^3)$, which is the ratio between $n(n-1)$ and n^2. This error factor becomes quite problematic because $\rho\ell_1^3 \sim |\log\rho|^{1+3\eta} \gg \epsilon$. To handle this problem, we merge 2^{3h} adjacent small cubic cells of side length ℓ_1 into a larger cubic box of side length ℓ_2. (Recall that we chose $\ell_2 = 2^h\ell_1$.) For a technical reason, we only merge two cells at a time, and we apply perturbation theory to achieve a similar result to Lemma 2.6 for this 'doubled box.' When we get to the cell of side length ℓ_2, we can obtain the following lemma:

Lemma 2.7. *Let $x_1, x_2, \cdots, x_N \in \Lambda$. Let $\Lambda_{\ell_2} \subset \Lambda$ be a cubic box of side length $\ell_2 \sim \rho^{-\frac{4}{9}}|\log\rho|^{-\frac{2}{3}}$. Divide Λ_{ℓ_2} into 2^{3h} smaller cubic cells of side length ℓ_1 such that $\ell_1 = 2^{-h}\ell_2$. Call those small* *cubic cells* $B_1, B_2, \cdots, B_{2^{3h}}$. *Define* f *as in* (2.26) *and let*

$$\widetilde{f}(t) := \begin{cases} t(t-1) & \text{if } t \le \rho\ell_2^3 \\ (2\rho\ell_2^3 - 1)t - (\rho\ell_2^3)^2 & \text{if } t > \rho\ell_2^3 \end{cases}. \tag{2.28}$$

Fix $x_{n+1}, x_{n+2}, \cdots, x_N$ *outside* Λ_{ℓ_2} *and let* $\boldsymbol{x}_n = (x_1, x_2, \cdots, x_n)$. *Define* $\mathcal{N}(B_k)$ *as*

$$\mathcal{N}(B_k) := \sum_{j=1}^{n} 1(x_j \in B_k). \tag{2.29}$$

Then, for any $\psi(\boldsymbol{x}_n) \in L^2(\Lambda_{\ell_2}^n)$ *with* $\phi_j = W_j^{-1}\psi$ *for* $j = 1, 2, \cdots, n$,

$$\begin{aligned} &\left(\frac{1}{2} - \epsilon\right) \sum_{j=1}^{n} \int_{\Lambda_{\ell_2}^n} |W_j|^2 |\nabla_j \phi_j(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n + \frac{4\pi a}{\ell_1^3} \sum_{k=1}^{2^{3h}} \int_{\Lambda_{\ell_2}^n} f(\mathcal{N}(B_k)) |\psi(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n \\ \ge\ & \widetilde{f}(n)(1 - C\rho^{\frac{1}{3}} |\log\rho|) \frac{4\pi a}{\ell_2^3} \|\psi\|_2^2. \end{aligned} \tag{2.30}$$

Proofs of the key lemmas above are given in Section 3 and Section 4.

2.4. **Proof of main theorem.**

Proof of Theorem 2.2. Assume that ρ is sufficiently small. Lemma 2.3 shows that, for any $\Psi \in L^2(\Lambda^N)$,

$$\langle \Psi, H_N \Psi \rangle \ge \epsilon \langle \Psi, H_N \Psi \rangle + (1-\epsilon)\Big(\sum_{j=1}^{N} \int_{\Lambda^N} |W_j(\mathbf{x})|^2 |\nabla_j \Phi_j(\mathbf{x})|^2 d\mathbf{x} + \sum_{i \ne j}^{N} \int_{\Lambda^N} q_{ij} |\Psi(\mathbf{x})|^2 d\mathbf{x} \Big), \tag{2.31}$$

where $\Phi_j = W_j^{-1}\Psi$ and $\epsilon = \rho^{\frac{1}{3}} |\log\rho|^3$.

Recall that in Lemma 2.5 we divided Λ into small cubic cells $\Lambda_{u\lambda}$ of side length ℓ_1, where u is the origin of division and λ is an index for those cells. From Lemma 2.5, we get

$$\begin{aligned} &\langle \Psi, H_N \Psi \rangle \\ \ge\ & \epsilon \langle \Psi, H_N \Psi \rangle + (1-\epsilon) \sum_{j=1}^{N} \int_{\Lambda^N} |W_j(\mathbf{x})|^2 |\nabla_j \Phi_j(\mathbf{x})|^2 d\mathbf{x} \\ & + (1-\epsilon) \inf_u \left(\frac{\ell_1}{\ell_1 - 4\ell_0}\right)^3 \sum_{i \ne j}^{N} \int_{\Lambda^{N-2}} dx_1 dx_2 \cdots \widehat{dx_i} \cdots \widehat{dx_j} \cdots dx_N \Big(\sum_{\lambda} \int_{\Lambda_{u\lambda}^2} \widetilde{q_{ij}} |\Psi(\mathbf{x})|^2 dx_i dx_j \Big). \end{aligned} \tag{2.32}$$

For a given box B, Let $\mathcal{N}(B)$ be the function that indicates the number of particles in B, i.e.,

$$\mathcal{N}(B) := \sum_{i=1}^{N} 1(x_i \in B). \tag{2.33}$$

Applying Lemma 2.6, we get

$$\begin{aligned} & \epsilon \langle \Psi, H_N \Psi \rangle + \frac{1}{2} \sum_{j=1}^{N} \int_{\Lambda^N} |W_j(\mathbf{x})|^2 |\nabla_j \Phi_j(\mathbf{x})|^2 d\mathbf{x} \\ & + (1-\epsilon) \left(\frac{\ell_1}{\ell_1 - 4\ell_0}\right)^3 \sum_{i \ne j}^{N} \int_{\Lambda^{N-2}} dx_1 dx_2 \cdots \widehat{dx_i} \cdots \widehat{dx_j} \cdots dx_N \Big(\sum_{\lambda} \int_{\Lambda_{u\lambda}^2} \widetilde{q_{ij}} |\Psi(\mathbf{x})|^2 dx_i dx_j \Big) \\ \ge\ & (1 - C\epsilon) \frac{4\pi a}{\ell_1^3} \int_{\Lambda^N} \sum_{\lambda} f(\mathcal{N}(\Lambda_{u\lambda})) |\Psi(\mathbf{x})|^2 d\mathbf{x}, \end{aligned} \tag{2.34}$$

where f is defined as in (2.26). Note that in (2.34) we only used half of kinetic energy. Combining (2.31), (2.32), and (2.34), we get

$$\langle \Psi, H_N \Psi \rangle \geq (\frac{1}{2} - \epsilon) \sum_{j=1}^{N} \int_{\Lambda^N} |W_j(\mathbf{x})|^2 |\nabla_j \Phi_j(\mathbf{x})|^2 d\mathbf{x} + (1 - C\epsilon) \frac{4\pi a}{\ell_1^3} \int_{\Lambda^N} \sum_{\lambda} f(\mathcal{N}(\Lambda_{u\lambda})) |\Psi(\mathbf{x})|^2 d\mathbf{x}. \tag{2.35}$$

Now we consider larger cubic boxes of side length ℓ_2. We let $\Lambda'_{u\theta}$ be a cubic box of side length ℓ_2, which is a union of 2^{3h} $\Lambda_{u\lambda}$'s, where θ is an index for those larger cubic boxes. With a shorthand notation $\mathbf{x} = (x_1, x_2, \cdots, x_N)$, we obtain from Lemma 2.7 that

$$\begin{aligned} & (\frac{1}{2} - \epsilon) \sum_{j=1}^{N} \int_{\Lambda^N} |W_j(\mathbf{x})|^2 |\nabla_j \Phi_j(\mathbf{x})|^2 d\mathbf{x} + (1 - C\epsilon) \frac{4\pi a}{\ell_1^3} \int_{\Lambda^N} \sum_{\lambda} f(\mathcal{N}(\Lambda_{u\lambda})) |\Psi(\mathbf{x})|^2 d\mathbf{x} \\ \geq \quad & (1 - C\epsilon) \frac{4\pi a}{\ell_2^3} \int_{\Lambda^N} \sum_{\theta} \widetilde{f}(\mathcal{N}(\Lambda'_{u\theta})) |\Psi(\mathbf{x})|^2 d\mathbf{x}, \end{aligned} \tag{2.36}$$

where $\widetilde{f}$ is a convex function such that $\widetilde{f}(n) = n(n-1)$ if $n \leq \rho \ell_2^3$. So far we have proved that

$$\langle \Psi, H_N \Psi \rangle \geq \inf_u \, (1 - C\epsilon) \frac{4\pi a}{\ell_2^3} \int_{\Lambda^N} \sum_{\theta} \widetilde{f}(\mathcal{N}(\Lambda'_{u\theta})) |\Psi(\mathbf{x})|^2 d\mathbf{x}. \tag{2.37}$$

We are left to minimize $\sum_\theta \widetilde{f}(\mathcal{N}(\Lambda'_{u\theta}))$. Since $\widetilde{f}$ is convex and $\sum_\theta \mathcal{N}(\Lambda'_{u\theta}) = N$, we can use Jensen's inequality to get, for any $x_1, x_2, \cdots, x_N$,

$$\sum_{\theta} \widetilde{f}(\mathcal{N}(\Lambda'_{u\theta})) \geq \frac{L^3}{\ell_2^3} \widetilde{f}(\rho \ell_2^3) \geq \frac{L^3}{\ell_2^3} \rho \ell_2^3 (\rho \ell_2^3 - 1) \geq (1 - C\rho^{\frac{1}{3}} |\log \rho|^2) N \rho \ell_2^3. \tag{2.38}$$

Note that this lower bound is independent of u. Hence

$$\frac{\langle \Psi, H_N \Psi \rangle}{\langle \Psi, \Psi \rangle} \geq (1 - C\epsilon) \frac{4\pi a}{\ell_2^3} N \rho \ell_2^3 = 4\pi a \rho N (1 - C\epsilon). \tag{2.39}$$

This shows that, when ρ is small enough, there exists a constant C_0 such that

$$\lim_{\substack{N, L \to \infty \\ N/L^3 = \rho}} \frac{E(N, \Lambda)}{N} \geq 4\pi a \rho (1 - C_0 \epsilon), \tag{2.40}$$

which was to be proved.

□

3. Lower bound estimates

In this section, we prove Lemma 2.3, Lemma 2.5, and Lemma 2.6, which were used in the proof of main theorem.

3.1. **Conversion into a soft potential.**

Proof of Lemma 2.3. Fix j and consider $x_1, x_2, \cdots, \widehat{x_j}, \cdots, x_N$ to be fixed. From that $\Phi_j = W_j^{-1} \Psi$, we get

$$\begin{aligned} & \int_{\Lambda} |\nabla_j \Psi(\mathbf{x})|^2 dx_j + \frac{1}{2} \sum_{i: i \neq j} \int_{\Lambda} V(x_i - x_j) |\Psi(\mathbf{x})|^2 dx_j \\ = \quad & \int_{\Lambda} |W_j(\mathbf{x})|^2 |\nabla_j \Phi_j(\mathbf{x})|^2 dx_j + \sum_{i: i \neq j} \int_{\Lambda} \Big[W_j(\mathbf{x}) \big(-\Delta_j + \frac{1}{2} V(x_i - x_j) \big) W_j(\mathbf{x}) \Big] |\Phi_j(\mathbf{x})|^2 dx_j \end{aligned} \tag{3.1}$$

For each $i \neq j$, we have either $F_{ij} = 1$ or $F_{ij} = 0$. Note that F_{ij} is independent of x_j.

(1) When $F_{ij}=1$, consider $B(x_i,\ell_0)$, a ball of radius ℓ_0 centered at x_i. When $x_j\in B(x_i,\ell_0)$, $W_j(\mathbf{x})=1-\tau(\ell_0,x_i-x_j)$, thus

$$\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})=e_0(\ell_0)W_j(\mathbf{x}). \tag{3.2}$$

When $x_j\notin B(x_i,\ell_0)$, $\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})=0$. Thus,

$$\begin{aligned}&\int_\Lambda\Big[W_j(\mathbf{x})\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})\Big]|\Phi_j(\mathbf{x})|^2dx_j\\ =\;&\int_\Lambda e_0(\ell_0)1(|x_i-x_j|\le\ell_0)|W_j(\mathbf{x})|^2|\Phi_j(\mathbf{x})|^2dx_j=\int_\Lambda F_{ij}(\mathbf{x})q(\ell_0,x_i-x_j)|\Psi(\mathbf{x})|^2dx_j.\end{aligned} \tag{3.3}$$

(2) When $F_{ij}=0,G_{ij}=1$, consider $B(x_i,t_{ij})$, a ball of radius t_{ij} centered at x_i. When $x_j\in B(x_i,t_{ij})$, $W_i(\mathbf{x})=1-\tau(t_{ij},x_i-x_j)$, thus

$$\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})=e_0(t_{ij})W_j(\mathbf{x}). \tag{3.4}$$

When $x_j\notin B(x_i,t_{ij})$, $W_j(\mathbf{x})\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})=0$. Thus,

$$\int_\Lambda\Big[W_j(\mathbf{x})\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})\Big]|\Phi_j(\mathbf{x})|^2dx_j \tag{3.5}$$
$$=\int_\Lambda e_0(t_{ij})1(|x_i-x_j|\le t_{ij})|W_j(\mathbf{x})|^2|\Phi_j(\mathbf{x})|^2dx_j=\int_\Lambda(1-F_{ij}(\mathbf{x}))G_{ij}(\mathbf{x})q(t_{ij},x_i-x_j)|\Psi(\mathbf{x})|^2dx_j.$$

(3) When $F_{ij}=G_{ij}=0$, we have that $W_j(\mathbf{x})=1$, thus we get

$$\int_\Lambda\Big[W_j(\mathbf{x})\big(-\Delta_j+\frac{1}{2}V(x_i-x_j)\big)W_j(\mathbf{x})\Big]|\Phi_j(\mathbf{x})|^2dx_j\ge0. \tag{3.6}$$

¿From cases (1)-(3), (3.1) implies

$$\begin{aligned}&\int_\Lambda\big(|\nabla_j\Psi(x_j)|^2+\frac{1}{2}\sum_{i:i\ne j}V(x_i-x_j)|\Psi(x_j)|^2\big)dx_j\\ \ge\;&\int_\Lambda|W_j(\mathbf{x})|^2|\nabla_j\Phi_j(\mathbf{x})|^2dx_j\\ &+\sum_{i:i\ne j}\int_\Lambda\big[F_{ij}(\mathbf{x})q(\ell_0,x_i-x_j)+(1-F_{ij}(\mathbf{x}))G_{ij}(\mathbf{x})q(t_{ij},x_i-x_j)\big]|\Psi(\mathbf{x})|^2dx_j.\end{aligned} \tag{3.7}$$

To get back to the N particle problem, we first integrate (3.7) and summing it over j gives the desired lemma. □

3.2. **Decomposition of Λ.**

Proof of Lemma 2.5. It suffices to prove the lemma for $u\in[-\ell_1/2,\ell_1/2)^3=\Gamma$. Let $G:=\ell_1\mathbb{Z}^3\cap\Lambda$. For $\lambda\in G$, we let $\Lambda_{u\lambda}$ be a cubic box of side length ℓ_1 centered at $(u+\lambda)$. Here, u corresponds to the origin of the grid that divides Λ into small boxes $\Lambda_{u\lambda}$. Note that the positions of those boxes depend on u.

Define a $\widetilde\chi_{u\lambda}$ by

$$\widetilde\chi(x):=\begin{cases}1 & \text{if } x\in[-\frac{\ell_1}{2}+2\ell_0,\frac{\ell_1}{2}-2\ell_0)^3\\ 0 & \text{otherwise}\end{cases}. \tag{3.8}$$

and $\widetilde\chi_{u\lambda}:=\widetilde\chi(x-u-\lambda)$. If $\widetilde\chi_{u\lambda}(x_i)=1$, then $x_i\in\Lambda_{u\lambda}$ and x_i is not within distance $2\ell_0$ from the boundary of $\Lambda_{u\lambda}$. Note that

$$\frac{1}{|\Gamma|}\int_\Gamma du\sum_{\lambda\in G}\widetilde\chi_{u\lambda}(x)=(\frac{\ell_1-4\ell_0}{\ell_1})^3 \tag{3.9}$$

for any $x \in \Lambda$. This means that the probability of having $x \in \Lambda$ to satisfy $\widetilde{\chi}_{u\lambda} = 1$ for a λ is $(\ell_1 - 4\ell_0)^3/\ell_1^3$.

¿From the definitions of $\widetilde{q_{ij}}$ and $\widetilde{\chi}_{u\lambda}$, we have

$$\widetilde{q_{ij}}(\mathbf{x}) = \sum_{\lambda \in G} \widetilde{\chi}_{u\lambda}(x_i) q_{ij}(\mathbf{x}). \tag{3.10}$$

($\widetilde{q_{ij}}$ depends on u, but we omit it.) Thus, for any $\Psi(\mathbf{x}) \in L^2(\Lambda^N)$,

$$\begin{aligned} \frac{1}{|\Gamma|}\int_\Gamma du \int_{\Lambda^2} \widetilde{q_{ij}} |\Psi(\mathbf{x})|^2 dx_i dx_j &= \frac{1}{|\Gamma|}\int_{\Lambda^2}\int_\Gamma du \sum_{\lambda \in G} \widetilde{\chi}_{u\lambda}(x_i) q_{ij} |\Psi(\mathbf{x})|^2 dx_i dx_j \\ = \quad (\frac{\ell_1 - 4\ell_0}{\ell_1})^3 \int_{\Lambda^2} q_{ij} |\Psi(\mathbf{x})|^2 dx_i dx_j. \end{aligned} \tag{3.11}$$

Integrating (3.11) with respect to $dx_1 dx_2 \cdots \widehat{dx_i} \cdots \widehat{dx_j} \cdots dx_N$ and summing over i, j gives

$$\begin{aligned} \sum_{i \neq j}^N \int_{\Lambda^N} q_{ij} |\Psi(\mathbf{x})|^2 d\mathbf{x} &= \frac{1}{|\Gamma|}\int_\Gamma du (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i\neq j}^N \int_{\Lambda^N} \widetilde{q_{ij}} |\Psi(\mathbf{x})|^2 d\mathbf{x} \\ \geq \quad \inf_u (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i\neq j}^N \int_{\Lambda^N} \widetilde{q_{ij}} |\Psi(\mathbf{x})|^2 d\mathbf{x}. \end{aligned} \tag{3.12}$$

This proves the lemma. □

3.3. **Lower bound estimate - Small cubic cell.**

Proof of Lemma 2.6. We consider the following cases:

(1) When $n \leq 9\rho\ell_1^3$:

When $n \leq 1$, $f(n) = 0$ and the lemma is trivial. Suppose that $n \geq 2$. Let

$$T_j = -W_j^{-1}\nabla_j W_j^2 \nabla_j W_j^{-1}, \tag{3.13}$$

defined on the functions in $L^2(B^n)$ with W_j-Neumann boundary conditions in B. For $i, j \in \{1, 2, \cdots, n\}$, $i \neq j$, we want to consider an operator

$$\frac{T_i}{4(n-1)} + \frac{T_j}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \widetilde{q_{ij}}, \tag{3.14}$$

which is defined on functions in $L^2(B^n)$ satisfying W_i-Neumann boundary conditions and W_j-Neumann boundary conditions.

We first estimate a lower bound for

$$\frac{T_j}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \widetilde{q_{ij}} \tag{3.15}$$

by Temple's inequality [18], with $[4(n-1)]^{-1}T_j$ as the unperturbed part in the first order perturbation theory.

To find the gap of T_j, we first notice that W_j, which is defined in (2.14), is the ground state of T_j. For a function $\psi(x_j)$ with W_j-Neumann boundary condition in B with

$$\int_B W_j \psi(x_j) dx_j = 0, \tag{3.16}$$

we apply Poincare's inequality to obtain

$$\int_B |\psi(x_j) - \bar{\psi}|^2 dx_j \leq C|B|^{\frac{2}{3}} \int_B |\nabla\psi(x_j)|^2 dx_j, \tag{3.17}$$

where

$$\bar{\psi} = \frac{1}{|B|}\int_B \psi(x_j) dx_j. \tag{3.18}$$

We also have

$$\begin{aligned}&\int_B |\psi(x_j)-\bar{\psi}|^2 dx_j = \int_B |\psi(x_j)|^2 dx_j - \frac{1}{|B|}\Big(\int_B \psi(x_j)dx_j\Big)^2\\
= &\int_B |\psi(x_j)|^2 dx_j - \frac{1}{|B|}\Big(\int_B (1-W_j)\psi(x_j)dx_j\Big)^2 \geq \int_B |\psi(x_j)|^2 dx_j - \frac{c_0^2}{|B|}\Big(\int_B |\psi(x_j)|dx_j\Big)^2\\
\geq &(1-c_0^2)\int_B |\psi(x_j)|^2 dx_j. \end{aligned} \tag{3.19}$$

Thus, we can see that, in a box of side length ℓ_1,

$$(\text{gap of } T_j) \geq C\ell_1^{-2}. \tag{3.20}$$

Now, we get from Temple's inequality that

$$\frac{T_j}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \widetilde{q_{ij}} \geq (\frac{\ell_1}{\ell_1-4\ell_0})^3 \langle \widetilde{q_{ij}} \rangle_{W_j} - C\frac{\langle \widetilde{q_{ij}}^2\rangle_{W_j} - \langle \widetilde{q_{ij}}\rangle_{W_j}^2}{n_\lambda^{-1}\ell_1^{-2}}, \tag{3.21}$$

where $\langle F\rangle_{W_i}$ denotes

$$\langle F\rangle_{W_i} = \int_B dx_i F|W_i|^2 / \int_B dx_i |W_i|^2. \tag{3.22}$$

Note that validity of (3.21) can be easily checked, since the gap of $[4(n-1)]^{-1}T_j$ is much larger than the expectation value of $\widetilde{q_{ij}}$ in the ground state W_j as we can see from

$$Cn^{-1}\ell_1^{-2} \geq C\rho^{\frac{2}{3}}|\log\rho|^{-\frac{5}{3}-5\eta} \gg C\rho^{\frac{2}{3}}|\log\rho|^{-2} \geq \|q_{ij}\|_\infty \geq \langle \widetilde{q_{ij}}\rangle_{W_j}. \tag{3.23}$$

We want to estimate $\langle \widetilde{q_{ij}}\rangle_{W_j}$ and $\langle \widetilde{q_{ij}}^2\rangle_{W_j}$. It follows from Lemma 5.1 that, for all $\ell_{-1}\leq \kappa \leq \ell_0$,

$$\tau(\kappa, x_i - x_j) \leq \frac{C}{|x_i - x_j|} \tag{3.24}$$

Hence,

$$\widetilde{q_{ij}}|W_j|^2 \geq \widetilde{q_{ij}}(1-\frac{C}{|x_i-x_j|})^2 \geq \widetilde{q_{ij}}(1-\frac{C}{|x_i-x_j|}). \tag{3.25}$$

Let S_{-1} and S_0 be sets of all points in B that are not within distance $2\ell_{-1}$ and $2\ell_0$, respectively, from any of $x_1, x_2, \cdots, \widehat{x_i}, \cdots, \widehat{x_j}, \cdots, x_N$, i.e.,

$$S_{-1} = \{x \in B : \forall k \neq i,j, |x - x_k| > 2\ell_{-1}\}, \tag{3.26}$$

$$S_0 = \{x \in B : \forall k \neq i,j, |x - x_k| > 2\ell_0\}, \tag{3.27}$$

and $\widetilde{B}$ be a set of all points in B that are not within distance $2\ell_0$ to the boundary of B, i.e.,

$$\widetilde{B} = \{x \in B : d(x, \partial B) \geq 2\ell_0\}. \tag{3.28}$$

By definition, $\widetilde{q_{ij}} = 0$ if $x_i \in B\backslash S_{-1}$ or $x_i \in B\backslash \widetilde{B}$, and $\widetilde{q_{ij}} = q_{ij}$ if $x_i \in S_{-1}\cap\widetilde{B}$.

When $x_i \in S_0 \cap \widetilde{B}$,

$$\begin{aligned}\int_B \widetilde{q_{ij}}|W_j|^2 dx_j &\geq \int_B \widetilde{q_{ij}}(1-\frac{C}{|x_i-x_j|})dx_j \geq e_0(\ell_0)\int_{|x_i-x_j|\leq\ell_0}(1-\frac{C}{|x_i-x_j|})dx_j\\
&\geq (1-\frac{C}{\ell_0})e_0(\ell_0)\int_{|x_i-x_j|\leq\ell_0} 1 dx_j \geq 4\pi a(1-\frac{C}{\ell_0}).\end{aligned} \tag{3.29}$$

When $x_i \in (S_{-1}\backslash S_0)\cap\widetilde{B}$,

$$\begin{aligned}\int_B \widetilde{q_{ij}}|W_j|^2 dx_j &\geq \int_B \widetilde{q_{ij}}(1-\frac{C}{|x_i-x_j|})dx_j \geq e_0(t_{ij})\int_{|x_i-x_j|\leq t_{ij}}(1-\frac{C}{|x_i-x_j|})dx_j\\
&\geq (1-\frac{C}{t_{ij}})e_0(t_{ij})\int_{|x_i-x_j|\leq t_{ij}} 1 dx_j \geq 4\pi a(1-\frac{C}{t_{ij}}).\end{aligned} \tag{3.30}$$

For $\langle \widetilde{q_{ij}}^2\rangle_{W_j}$, when $x_i \in S_0 \cap \widetilde{B}$,

$$\int_B \widetilde{q_{ij}}^2 |W_j|^2 dx_j \le \int_B \widetilde{q_{ij}}^2 dx_j \le C\ell_0^{-3}, \tag{3.31}$$

and, when $x_i \in (S_{-1}\backslash S_0) \cap \widetilde{B}$,

$$\int_B \widetilde{q_{ij}}^2 |W_j|^2 dx_j \le \int_B \widetilde{q_{ij}}^2 dx_j \le Ct_{ij}^{-3}. \tag{3.32}$$

Since we know from Lemma 5.2 that

$$\ell^3(1 - Cn\frac{\ell_0^2}{\ell^3} - C\ell^{-1}) \le \int_{\Lambda_\ell} |W_j|^2 dx_j \le \ell^3, \tag{3.33}$$

inserting (3.29), (3.30), (3.31), (3.32), and (3.33) into (3.21), we obtain that

$$\begin{aligned}
&\frac{T_j}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \widetilde{q_{ij}} \\
\ge\ & (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \frac{4\pi a}{\ell_1^3} \left((1 - \frac{C}{\ell_0} - C\frac{n\ell_1^2}{\ell_0^3}) \cdot 1(x_i \in S_0 \cap \widetilde{B}) + (1 - \frac{C}{t_{ij}} - C\frac{n\ell_1^2}{t_{ij}^3}) \cdot 1(x_i \in (S_{-1}\backslash S_0) \cap \widetilde{B}) \right) \\
\ge\ & (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \frac{4\pi a}{\ell_1^3} \left((1 - C\frac{n\ell_1^2}{\ell_0^3}) \cdot 1(x_i \in S_0 \cap \widetilde{B}) + (1 - C\frac{n\ell_1^2}{t_{ij}^3}) \cdot 1(x_i \in (S_{-1}\backslash S_0) \cap \widetilde{B}) \right).
\end{aligned} \tag{3.34}$$

Here, for the last inequality, we used $n\ell_1^2 \gg \ell_0^2$ and, when $x_i \in (S_{-1}\backslash S_0) \cap \widetilde{B}$, $n\ell_1^2 \gg t_{ij}^2$.

Let

$$\xi_i(\mathbf{x}_n) = \frac{4\pi a}{\ell_1^3} \left((1 - C\frac{n\ell_1^2}{\ell_0^3}) \cdot 1(x_i \in S_0 \cap \widetilde{B}) + (1 - C\frac{n\ell_1^2}{t_{ij}^3}) \cdot 1(x_i \in (S_{-1}\backslash S_0) \cap \widetilde{B}) \right). \tag{3.35}$$

Note that S_0 and S_{-1} are independent of x_i and x_j, and ξ_i is independent of x_j. We apply Temple's inequality to

$$\frac{T_i}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \xi_i \tag{3.36}$$

with $[4(n-1)]^{-1}T_i$ as the unperturbed part. Then, we get

$$\frac{T_i}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \xi_i \ge (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \langle \xi_i \rangle_{W_i} - C\frac{\langle \xi_i^2 \rangle_{W_i}}{n^{-1}\ell_1^{-2}}. \tag{3.37}$$

We are to estimate $\langle \xi_i \rangle_{W_i}$. By definition,

$$\begin{aligned}
&\langle \xi_i \rangle_{W_i} \ge \ell_1^{-3} \int_{S_{-1}\cap\widetilde{B}} \xi_i |W_i|^2 dx_i \\
\ge\ & \frac{4\pi a}{\ell_1^3}(1 - C\frac{n\ell_1^2}{\ell_0^3}) \int_{S_{-1}\cap\widetilde{B}} \ell_1^{-3}|W_i|^2 dx_i - \frac{C}{\ell_1^3} \int_{(S_{-1}\backslash S_0)\cap\widetilde{B}} \ell_1^{-3} \frac{n\ell_1^2}{t_{ij}^3} dx_i.
\end{aligned} \tag{3.38}$$

To estimate the last term in the right hand side, we note that (1) $t_{ij} = \frac{1}{2}|x_i - x_k|$ for some x_k other than x_j and (2) the union of annuli $\{x : 2\ell_{-1} \le |x - x_k| \le 2\ell_0\}$ for all $x_k \in B$ other than x_i and x_j covers $(S_{-1}\backslash S_0)$. Thus,

$$\int_{(S_{-1}\backslash S_0)\cap\widetilde{B}} t_{ij}^{-3} dx_i \le \sum_{k:k\ne i} \int_{2\ell_{-1} \le |x_i - x_k| \le 2\ell_0} |x_i - x_k|^{-3} dx_i. \tag{3.39}$$

To estimate the first term in the right hand side of (3.38), we use

$$\int_{S_{-1}\cap\widetilde{B}} |W_i|^2 dx_i \ge \int_{S_{-1}\cap\widetilde{B}} dx_i - \int_B (1 - |W_i|^2) dx_i \ge |S_{-1} \cap \widetilde{B}| - Cn\ell_0^2, \tag{3.40}$$

where the last inequality follows from Lemma 5.2. Since $|S_{-1}| \geq \ell_1^3 - Cn(\ell_{-1})^3$ and $|\widetilde{B}| = (\ell_1 - 4\ell_0)^3$, from (3.38), (3.39), and (3.40), we get

$$\begin{aligned}
&\langle \xi_i \rangle_{W_i} \\
\geq\ & \frac{|S_{-1} \cap \widetilde{B}|}{\ell_1^3} \frac{4\pi a}{\ell_1^3} (1 - C\frac{n\ell_1^2}{\ell_0^3})(1 - Cn\frac{\ell_0^2}{\ell_1^3}) - Cn \sum_{k:k\neq i} \int_{2\ell_{-1} \leq |x_i - x_k| \leq 2\ell_0} \ell_1^{-4} |x_i - x_k|^{-3} dx_i \\
\geq\ & (\frac{\ell_1 - 4\ell_0}{\ell_1})^3 (1 - Cn\frac{(\ell_{-1})^3}{\ell_1^3}) \frac{4\pi a}{\ell_1^3} (1 - C\frac{n\ell_1^2}{\ell_0^3})(1 - Cn\frac{\ell_0^2}{\ell_1^3})(1 - Cn^2 \frac{|\log \rho|}{\ell_1}). \qquad (3.41)
\end{aligned}$$

Since $\langle \xi_i^2 \rangle_{W_i} \leq \|\xi_i\|_\infty^2 \leq C\ell_1^{-6}$,

$$\frac{\langle \xi_i^2 \rangle_{W_i}}{n^{-1}\ell_1^{-2}} \leq Cn\ell_1^{-4}. \qquad (3.42)$$

Thus, combining (3.34), (3.37), (3.41), and (3.42), we obtain that

$$\begin{aligned}
&\frac{T_i}{4(n-1)} + \frac{T_j}{4(n-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \widetilde{q_{ij}} \\
\geq\ & (1 - Cn\frac{(\ell_{-1})^3}{\ell_1^3})(1 - C\frac{n\ell_1^2}{\ell_0^3})(1 - Cn\frac{\ell_0^2}{\ell_1^3})(1 - Cn^2 \frac{|\log \rho|}{\ell_1}) \frac{4\pi a}{\ell_1^3} \qquad (3.43) \\
\geq\ & (1 - C\rho^{\frac{1}{3}} |\log \rho|^3) \frac{4\pi a}{\ell_1^3}.
\end{aligned}$$

This implies, when $x_{n+1}, \cdots, x_N$ are outside of B, for any $\psi(\mathbf{x}_n) \in L^2(B^n)$ with $\phi_l = W_l^{-1}\psi$ for $l = 1, 2, \cdots, n$,

$$\begin{aligned}
&\frac{1}{4(n-1)} \int_{B^n} \big(|W_i|^2 |\nabla_i \phi_i(\mathbf{x}_n)|^2 + |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2\big) d\mathbf{x}_n + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \int_{B^n} \widetilde{q_{ij}} |\psi(\mathbf{x}_n)|^2 d\mathbf{x}_n \\
\geq\ & (1 - C\epsilon) \frac{4\pi a}{\ell_1^3} \|\psi\|_2^2. \qquad (3.44)
\end{aligned}$$

This shows how we can get the lower bound for the ground state energy in a soft potential regime. $\widetilde{q_{ij}}$ depends on particles other than x_i and x_j, but overall effect from them is insignificant and becomes a small error.

We apply this argument to all $1 \leq i, j \leq n$, $i \neq j$. After summing over all indices i and j, we get

$$\begin{aligned}
&\frac{1}{2} \sum_{j=1}^n \int_{B^n} |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i\neq j}^n \int_{B^n} \widetilde{q_{ij}} |\psi(\mathbf{x}_n)|^2 d\mathbf{x}_n \\
\geq\ & n(n-1)(1 - C\epsilon) \frac{4\pi a}{\ell_1^3} \|\psi\|_2^2. \qquad (3.45)
\end{aligned}$$

(2) When $9\rho\ell_1^3 < n \leq 9\epsilon^{-1}\rho\ell_1^3$:

Let $p := 9\rho\ell_1^3$. Here, n satisfies that

$$n\frac{(\ell_{-1})^3}{\ell_1^3} \ll 1, \quad np\frac{|\log \rho|}{\ell_1} \ll 1, n\frac{\ell_0^2}{\ell_1^3} \ll 1. \qquad (3.46)$$

We form particle groups in B, each of which contains p particles. This generates $\lfloor n/p \rfloor < n/2p$ groups and we call those groups $G_1, G_2, \cdots, G_{\lfloor n/p \rfloor}$.

For $i, j \in G_k$, we consider

$$\frac{T_i}{4(p-1)} + \frac{T_j}{4(p-1)} + (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \widetilde{q_{ij}} \qquad (3.47)$$

which is defined on functions in $L^2(B^n)$ satisfying W_i-Neumann boundary conditions and W_j-Neumann boundary conditions. We then use the Temple's inequality as in case (1) to get, for any $\psi(\mathbf{x}_n) \in L^2(B^n)$ with $\phi_l = W_l^{-1}\psi$ for $l = 1, 2, \cdots, n$,

$$\begin{aligned} & \frac{1}{4(n-1)} \int_{B^n} \left(|W_i|^2 |\nabla_i \phi_i(\mathbf{x}_n)|^2 + |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2\right) d\mathbf{x}_n + \left(\frac{\ell_1}{\ell_1 - 4\ell_0}\right)^3 \int_{B^n} \widetilde{q_{ij}} |\psi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n \\ \geq \; & (1 - C\frac{p\ell_1^2}{\ell_0^3})(1 - Cn\frac{\ell_0^2}{\ell_1^3})(1 - Cn\frac{(\ell_{-1})^3}{\ell_1^3})(1 - Cnp\frac{|\log\rho|}{\ell_1})\frac{4\pi a}{\ell_1^3}\|\psi\|_2^2. \end{aligned} \tag{3.48}$$

We apply this inequality to all particles in the particle group G_l, then to all particle groups in B. Using (3.46), we obtain

$$\begin{aligned} & \frac{1}{2}\sum_{j=1}^{n} \int_{B^n} |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n + \left(\frac{\ell_1}{\ell_1 - 4\ell_0}\right)^3 \int_{B^n} \widetilde{q_{ij}} |\psi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n \\ \geq \; & \lfloor \frac{n}{p} \rfloor p(p-1)(1 - Cn\frac{(\ell_{-1})^3}{\ell_1^3})(1 - C\frac{p\ell_1^2}{\ell_0^3})(1 - Cn\frac{\ell_0^2}{\ell_1^3})(1 - Cnp\frac{|\log\rho|}{\ell_1})\frac{4\pi a}{\ell_1^3}\|\psi\|_2^2 \qquad (3.49) \\ \geq \; & \frac{n}{2}(p-1)\left(\frac{8}{9}\right)\frac{4\pi a}{\ell_1^3}\|\psi\|_2^2 \geq (4\rho\ell_1^3 - 1)n(1 - C\epsilon)\frac{4\pi a}{\ell_1^3}\|\psi\|_2^2. \end{aligned}$$

(3) When $n > 9\epsilon^{-1}\rho\ell_1^3$:

In this case, we only use the term $\epsilon E(n, B)$ to prove the lemma and ignore the other terms, since they are non-negative. It is known that the ground state energy of n particle system in a box of side length ℓ_1,

$$E(n, B) \geq 4\pi a\left(\frac{n}{\ell_1^3}\right)n\left(1 - C\left(\frac{n}{\ell_1^3}\right)^{\frac{1}{17}}\right) \tag{3.50}$$

when the density (n/ℓ_1^3) is sufficiently small and

$$\ell_1 \geq C\left(\frac{n}{\ell_1^3}\right)^{-\frac{6}{17}}. \tag{3.51}$$

(See Theorem 2.4 in [14].)

In our case, if the box B contains $m := 9\epsilon^{-1}\rho\ell_1^3$ particles, then

$$\ell_1 \geq C\rho^{-\frac{1}{3}}|\log\rho|^{\frac{1}{3}} \gg C\rho^{-\frac{4}{17}}|\log\rho|^{\frac{18}{17}} = C\left(\frac{m}{\ell_1^3}\right)^{-\frac{6}{17}}, \tag{3.52}$$

and the density in this case

$$\frac{m}{\ell_1^3} = 9\epsilon^{-1}\rho = 9\rho^{\frac{2}{3}}|\log\rho|^{-3} \to 0 \tag{3.53}$$

as $\rho \to 0$. Thus we can indeed use (3.50) to obtain that

$$E(m, B) \geq 4\pi a\frac{m^2}{\ell_1^3}\left(1 - C\left(\frac{2m}{\ell_1^3}\right)^{\frac{1}{17}}\right) \tag{3.54}$$

To find a lower bound of $E(n, B)$, we form particle groups in B, each of which contains m particles. Since we have $\lfloor n/m \rfloor$ groups of size m, by superadditivity,

$$E(n, B) \geq \lfloor \frac{n}{m} \rfloor E(m, B) \geq (4\rho\ell_1^3)n(1 - C\epsilon)\frac{4\pi a}{\ell_1^3}. \tag{3.55}$$

Since, n falls into one of the above categories, we get the desired lemma. □

4. Box doubling method

In this section, we prove Lemma 2.7.

4.1. **Lower bound estimate - Large cubic cell.** In order to show Lemma 2.7, we need to prove a result analogous to Lemma 2.6 when Λ_{ℓ_2}, a box of side length ℓ_2 is given. We note:

$$\ell_1 \ll \ell_2 \sim \rho^{-\frac{4}{9}}|\log\rho|^{-\frac{2}{3}} \tag{4.1}$$

More specifically, for a box B with a side length between ℓ_1 and ℓ_2, we will show that the energy in B with n particles ($n \sim \rho|B|$) retains the form,

$$4\pi a(1-C\epsilon)|B|^{-1}n(n-1). \tag{4.2}$$

We can see that the error factor $1/n$ that comes from the ratio between $n(n-1)$ and n^2 becomes smaller as n increases, and eventually it becomes $1/n \sim \rho^{\frac{1}{3}}|\log\rho|^2 \ll \epsilon$ when the side length of B becomes ℓ_2, i.e, $n \sim \rho\ell_2^3$.

To demonstrate how to enlarge the size of box while retaining the form (4.2), we first consider a simple case where we have only two adjacent boxes Λ_A and Λ_B with the same size. Suppose that there are n particles in $\Lambda_A \cup \Lambda_B$. Let n_A and n_B denote the number of the particles in Λ_A and Λ_B. We assume the potential energy in Λ_A and Λ_B as $n_A(n_A-1)/|\Lambda_A|$ and $n_B(n_B-1)/|\Lambda_B|$, which has the form in (4.2). For $\alpha \in \mathbb{R}$, we define the Hamiltonian as

$$H_\alpha = \alpha\sum_{i=1}^{n}(-\Delta_i) + n_A(n_A-1)/|\Lambda_A| + n_B(n_B-1)/|\Lambda_B| \tag{4.3}$$

When $\alpha = 0$(no kinetic energy), the ground state energy of this Hamiltonian is $n(n-2)/|\Lambda_A \cup \Lambda_B|$. But when $\alpha = \infty$(particles are uniformly distributed in $\Lambda_A \cup \Lambda_B$), the ground state energy is equal to $n(n-1)/|\Lambda_A \cup \Lambda_B|$, which gives the desired form (4.2).

We will show that instead of $-\alpha\Delta_i$, a small potion of T_i can also guarantee the almost-uniform distribution and the desired form (4.2).

This heuristic argument shows our basic strategy in this section, which we call 'box doubling method.' Recall that $\ell_2 = 2^h\ell_1$. In this method, we begin from the first step where we have 2^{3h} small cubic cells of side length ℓ_1. We consider $(2^{3h}/2)$ pairs of adjacent boxes, and for each pair that consists of two adjacent boxes of same size Λ_A and Λ_B. As explained above, we can get a lower bound for the energy of n particle system in $\Lambda_A \cup \Lambda_B$ at expense of small potion of T_i's. In this way, we can effectively make the size of each box doubled, since the new 'potential energy term' in $\Lambda_A \cup \Lambda_B$ also has the form in (4.2)(when the density in $\Lambda_A \cup \Lambda_B$ is about ρ).

At the end of the first step, or the beginning of the 2nd step, we have 2^{3h-1} boxes whose dimensions are $\ell_1 \times \ell_1 \times 2\ell_1$. In the 2nd step, we consider 2^{3h-2} pairs of those boxes and perform the above process again for all the pairs. Keep using this method. At the beginning of the s-th step, we have 2^{3h-s} boxes, and after applying the above method to 2^{3h-s-1} pairs of boxes, the number of boxes gets halved and the size of each box doubled. And the new 'potential energy terms' in new boxes also have the form in (4.2)(when the density in new boxes is about ρ)

We keep repeating it until the side length of a box becomes ℓ_2, which is when $s = 3h$, and we only have one box left. The form of the potential term, (4.2) remains the same throughout this procedure, and it can lead us to the desired result, Lemma 2.7. We will make this argument rigorous in this section.

Before we begin the proof, we introduce definitions that will be used throughout this section.

- s is a non-negative integer that satisfies $1 \le s \le 3h$, where $\ell_2 = 2^h\ell_1$. We let

$$\ell(s) := 2^{\lfloor\frac{s-1}{3}\rfloor}\ell_1. \tag{4.4}$$

  $\ell(s)$ satisfies $\rho^{-\frac{1}{3}}|\log\rho|^{\frac{1}{3}+\eta} \sim \ell_1 \le \ell(s) \le \ell_2 \sim \rho^{-\frac{4}{9}}|\log\rho|^{-\frac{2}{3}}$, where $0 < \eta < 1/15$.

This s is a label keeping track of which step we are at. We begin from $s = 1$ and our method ends when $s = 3h$.

- Λ_A and Λ_B are boxes such that the volume of each box $|\Lambda_A| = |\Lambda_B| = 2^{s-1}\ell_1^3$ and the dimensions of Λ_A, Λ_B, and $\Lambda_A \cup \Lambda_B$ are either

(1) $\Lambda_A = \Lambda_B = \ell(s) \times \ell(s) \times \ell(s)$, $\Lambda_A \cup \Lambda_B = \ell(s) \times \ell(s) \times 2\ell(s)$,
(2) $\Lambda_A = \Lambda_B = \ell(s) \times \ell(s) \times 2\ell(s)$, $\Lambda_A \cup \Lambda_B = \ell(s) \times 2\ell(s) \times 2\ell(s)$, or
(3) $\Lambda_A = \Lambda_B = \ell(s) \times 2\ell(s) \times 2\ell(s)$, $\Lambda_A \cup \Lambda_B = 2\ell(s) \times 2\ell(s) \times 2\ell(s)$.

- $\mathcal{M}(A)$ and $\mathcal{M}(B)$ are the functions that indicate how many particles among $x_1, x_2, \cdots, x_n$ are in Λ_A and Λ_B, respectively, when $x_{n+1}, x_{n+2}, \cdots, x_N$ are outside $(\Lambda_A \cup \Lambda_B)$, i.e.,

$$\mathcal{M}(A) := \sum_{i=1}^{n} 1(x_i \in \Lambda_A), \quad \mathcal{M}(B) := \sum_{i=1}^{n} 1(x_i \in \Lambda_B) \tag{4.5}$$

Note that $\mathcal{M}(A)$ and $\mathcal{M}(B)$ depend on n though we omitted it.

Λ_A are Λ_B are a pair of boxes at the s-th step. Though we consider only two boxes at a time, note that we have 2^{3h-s-1} such pairs of boxes in s-th step.

Note that the potential term in Lemma 2.6, i.e., the right hand side of (2.27) is $4\pi a|B|^{-1}f(n)$, which is different from (4.2). $f(n)$ changes from quadratic to linear at $n = 2\rho\ell_1^3$. At the s-th step, we use f_s instead of f, and $f_s(n)$ becomes linear when $n \geq K_s$, i.e.,

- Define $(1 \leq s \leq 3h+1)$

$$f_s(t) := \begin{cases} t(t-1) & \text{if } t \leq K_s \\ (2K_s - 1)t - K_s^2 & \text{if } t > K_s \end{cases}. \tag{4.6}$$

The definition of f_s ensures that f_s is continuous and convex. And we define the parameters K_s as follows,

- Let

$$K_1 := 2\rho\ell_1^3 \tag{4.7}$$

and we choose K_s such that

$$2K_s - K_{s+1} \gg |\log\rho|^{\frac{1}{2}}\sqrt{K_{s+1}} \ \text{ and } \ K_s > 2^{s-1}\rho\ell_1^3. \tag{4.8}$$

For example,

$$K_s = \left(2 - (1 - \frac{1}{2^{\frac{s-1}{2}}})|\log\rho|^{-\eta}\right) \cdot 2^{s-1}\rho\ell_1^3. \tag{4.9}$$

We note that f_s with a suitable coefficient is our actual potential energy term in (4.2). When $s = 1$, f_s is equal to f in (2.26), and $f_s(t) = t(t-1)$ when $t \leq \rho \cdot 2^s\rho\ell_1^3$, i.e., the density is no more than ρ. (Note that $2^{s-1}\rho\ell_1^3$ is the volume of each box in the s-th step.)

Our proof of Lemma 2.7 requires the following proposition only, where we consider s to be fixed:

Proposition 4.1. *Let n be an integer and $1 \leq n \leq N$. Assume that $x_{n+1}, x_{n+2}, \cdots, x_N$ are fixed outside $(\Lambda_A \cup \Lambda_B)$. Then, for any $\psi(\boldsymbol{x}_n) \in L^2((\Lambda_A \cup \Lambda_B)^n)$ with $\phi_j = W_j^{-1}\psi$ for $j = 1, 2, \cdots, n$ (Here $\boldsymbol{x}_n = (x_1, x_2, \cdots, x_n)$),*

$$\begin{aligned} &\frac{(4\pi a)^{-1}}{|\log\rho|}\sum_{j=1}^{n}\int_{(\Lambda_A\cup\Lambda_B)^n}|W_j|^2|\nabla_j\phi_j(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n + \int_{(\Lambda_A\cup\Lambda_B)^n}\frac{f_s(\mathcal{M}(A)) + f_s(\mathcal{M}(B))}{|\Lambda_A|}|\psi(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n \\ \geq\ &(1 - C\rho^{\frac{1}{3}})\int_{(\Lambda_A\cup\Lambda_B)^n}\frac{f_{s+1}(n)}{|\Lambda_A \cup \Lambda_B|}|\psi(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n. \end{aligned} \tag{4.10}$$

Here, the constant C does not depend on s.

This proposition shows the outcome of box doubling method when it is applied to Λ_A and Λ_B at the s-th step. We can prove Lemma 2.7 from this proposition.

Proof of Lemma 2.7. Recall that we defined h as $\ell_2 = 2^h\ell_1$. Recall also that we have a cubic box Λ_{ℓ_2} whose side length is ℓ_2 and $B_1, B_2, \cdots, B_{2^{3h}}$ are small cubic cells of side length ℓ_1 such that $\bigcup_{k=1}^{2^{3h}} B_k = \Lambda_{\ell_2}$. $x_{n+1}, x_{n+2}, \cdots, x_N$ are fixed outside Λ_{ℓ_2} and we let $\mathbf{x}_n = (x_1, x_2, \cdots, x_n)$. Here, this n is different from the n in Proposition 4.1. We want to prove that, for a given $\psi(\mathbf{x}_n) \in L^2(\Lambda^n_{\ell_2})$ with $\phi_j = W_j^{-1}\psi$ for $j = 1, 2, \cdots, n$,

$$\begin{aligned}&(\frac{1}{2}-\epsilon)\sum_{j=1}^n \int_{\Lambda^n_{\ell_2}} |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n + \frac{4\pi a}{\ell_1^3}\sum_{k=1}^{2^{3h}} \int_{\Lambda^n_{\ell_2}} f(\mathcal{N}(B_k))|\psi(\mathbf{x}_n)|^2 d\mathbf{x}_n \\ \geq\quad & \widetilde{f}(n)(1-C\rho^{\frac{1}{3}}|\log\rho|)\frac{4\pi a}{\ell_2^3}\|\psi\|_2^2. \end{aligned} \tag{4.11}$$

Here,

$$\widetilde{f}(t) := \begin{cases} t(t-1) & \text{if } t \leq \rho\ell_2^3 \\ (2\rho\ell_2^3 - 1)t - (\rho\ell_2^3)^2 & \text{if } t > \rho\ell_2^3 \end{cases}. \tag{4.12}$$

Since

$$3h = \frac{\log(\ell_2^3/\ell_1^3)}{\log 2} \leq (\frac{1}{2}-\epsilon)|\log\rho|, \tag{4.13}$$

we can keep using Proposition 4.1 until we have only one box and its side length is ℓ_2. Then, we get

$$\begin{aligned}&(\frac{1}{2}-\epsilon)\sum_{j=1}^n \int_{\Lambda_{\ell_2}} |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n + 4\pi a\sum_{k=1}^{2^{3h}} \int_{\Lambda^n_{\ell_2}} \frac{f(\mathcal{N}(B_k))}{\ell_1^3}|\psi(\mathbf{x}_n)|^2 d\mathbf{x}_n \\ \geq\quad & (1-C\rho^{\frac{1}{3}}|\log\rho|)4\pi a\int_{\Lambda^n_{\ell_2}} \frac{f_{3h+1}(n)}{\ell_2^3}|\psi(\mathbf{x}_n)|^2 d\mathbf{x}_n. \end{aligned} \tag{4.14}$$

By definition (4.8), $K_{3h+1} > 2^{3h}\rho\ell_1^3 = \rho\ell_2^3$. Together with the definition of f in (2.26), it can be easily checked that $f_{3h+1} \geq \widetilde{f}$. This proves the desired result, (2.30). □

4.2. **Proof of Proposition 4.1.** To prove Proposition 4.1, we consider large n case and small n case separately. When n is large, we use the following lemma to prove Proposition 4.1.

Lemma 4.2. *Suppose that $n \geq K_{s+1} + 2\sqrt{K_{s+1}}$. Let $m_A, m_B = 0, 1, 2, \cdots, n$. Then,*

$$\min_{m_A+m_B=n} \frac{f_s(m_A) + f_s(m_B)}{|\Lambda_A|} \geq \frac{f_{s+1}(n)}{|\Lambda_A \cup \Lambda_B|}. \tag{4.15}$$

We note this inequality implies (4.10) directly, since the kinetic energy part is always non-negative. A proof of Lemma 4.2 will be given in the next subsection.

On the other hand, for small n case, i.e., when $n < K_{s+1} + 2\sqrt{K_{s+1}}$, we are going to prove the following inequality:

$$\begin{aligned}&\frac{(4\pi a)^{-1}}{|\log\rho|}\sum_{j=1}^n \int_{(\Lambda_A\cup\Lambda_B)^n} |W_j|^2 |\nabla_j \phi_j(\mathbf{x}_n)|^2 d\mathbf{x}_n + \int_{(\Lambda_A\cup\Lambda_B)^n} \frac{f_s(\mathcal{M}(A)) + f_s(\mathcal{M}(B))}{|\Lambda_A|}|\psi(\mathbf{x}_n)|^2 d\mathbf{x}_n \\ \geq\quad & (1-C\rho^{\frac{1}{3}})F_s(0,0,n)\|\psi\|_2^2 \\ \geq\quad & (1-C\rho^{\frac{1}{3}})\frac{f_{s+1}(n)}{|\Lambda_A\cup\Lambda_B|}\|\psi\|_2^2. \end{aligned} \tag{4.16}$$

Here, F_s is a function defined as follows:

Definition 4.3. For non-negative integers n_A, n_B, define

$$F_s(n_A, n_B, 0) := \frac{f_s(n_A)}{|\Lambda_A|} + \frac{f_s(n_B)}{|\Lambda_B|}. \tag{4.17}$$

To define $F(n_A, n_B, k)$, we use the following process: For a fixed k, consider k particles, $y_1, y_2, \cdots, y_k$, which are uniformly distributed in $\Lambda_A \cup \Lambda_B$, i.e.,

$$P(y_i \in \Lambda_A) = P(y_i \in \Lambda_B) = \frac{1}{2}. \tag{4.18}$$

Let $m_A(k)$ and $m_B(k)$ be the number of y_i's in Λ_A and Λ_B, respectively, i.e.,

$$m_A(k) = \sum_{i=1}^{k} 1(y_i \in \Lambda_A), \quad m_B(k) = \sum_{i=1}^{k} 1(y_i \in \Lambda_B) \tag{4.19}$$

Extend the definition of F so that

$$F_s(n_A, n_B, k) := \langle F_s(n_A + m_A(k), n_B + m_B(k), 0)\rangle_k, \tag{4.20}$$

where $\langle\cdot\rangle_k$ denotes expectation with respect to the distribution of $y_1, y_2, \cdots, y_k$. (We call those imaginary particles $y_1, y_2, \cdots, y_k$ 'randomized.')

To prove (4.16), we need the following lemmas that will be proved in subsection 4.3:

Lemma 4.4. *Suppose that* $n < K_{s+1} + 2\sqrt{K_{s+1}}$. *Then,*

$$F_s(0, 0, n) \geq (1-\rho)\frac{f_{s+1}(n)}{|\Lambda_A \cup \Lambda_B|}. \tag{4.21}$$

We note this lemma implies the second inequality of (4.16). For the first one we have:

Lemma 4.5. *Suppose that* $n < K_{s+1} + 2\sqrt{K_{s+1}}$. *Fix* $x_{n+1}, x_{n+2}, \cdots, x_N$ *outside* $(\Lambda_A \cup \Lambda_B)$ *and let* $\boldsymbol{x}_n = (x_1, x_2, \cdots, x_n)$. *Then, for any* $\psi(\boldsymbol{x}_n) \in L^2((\Lambda_A \cup \Lambda_B)^n)$ *with* $\phi_j = W_j^{-1}\psi$ *for* $j = 1, 2, \cdots, n$,

$$\begin{aligned}&\frac{(4\pi a)^{-1}}{|\log\rho|}\sum_{j=1}^{n}\int_{(\Lambda_A\cup\Lambda_B)^n} |W_j|^2|\nabla_j\phi_j(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n + \int_{(\Lambda_A\cup\Lambda_B)^n} F_s(\mathcal{M}(A), \mathcal{M}(B), 0)|\psi(\boldsymbol{x}_n)|^2 d\boldsymbol{x}_n \\ \geq\ & (1 - C\rho^{\frac{1}{3}})F_s(0,0,n)\|\psi\|_2^2.\end{aligned} \tag{4.22}$$

Now we are ready to prove Proposition 4.1.

Proof of Proposition 4.1. When $n \geq K_{s+1} + 2\sqrt{K_{s+1}}$, the desired result (4.10) follows from Lemma 4.2. When $n < K_{s+1} + 2\sqrt{K_{s+1}}$, we obtain (4.16) from Lemma 4.5 and Lemma 4.4, which implies (4.10). □

4.3. Proof of Lemma 4.2 and Lemma 4.4.

Proof of Lemma 4.2. Since f_s is convex,

$$\min_{m_A+m_B=n} f_s(m_A) + f_s(m_B) \geq 2f_s(\frac{n}{2}). \tag{4.23}$$

Thus, it suffices to prove

$$2f_s(\frac{n}{2}) \geq \frac{1}{2}f_{s+1}(n). \tag{4.24}$$

(1) When $K_{s+1} + 2\sqrt{K_{s+1}} \leq n \leq 2K_s$, we have

$$\begin{aligned}4f_s(\frac{n}{2}) - f_{s+1}(n) &= (n^2 - 2n) - \big[(2K_{s+1} - 1)n - K_{s+1}^2\big] \\ &= \big(n - K_{s+1} - \frac{1}{2}\big)^2 - K_{s+1} - \frac{1}{4}\end{aligned} \tag{4.25}$$

With $n - K_{s+1} \geq 2\sqrt{K_{s+1}}$, we have it is above zero.

(2) When $n > 2K_s$: we compare the derivatives of both sides of (4.24),

$$2\frac{d}{dn}\big(f_s(\frac{n}{2})\big) = 2K_s - 1 \geq K_{s+1} - \frac{1}{2} = \frac{1}{2}\frac{d}{dn}\big(f_{s+1}(n)\big). \tag{4.26}$$

Since we also have

$$2f_s(K_s) > \frac{1}{2}f_{s+1}(2K_s), \tag{4.27}$$

We can see that, for any $n > 2K_s$,

$$2f_s(\frac{n}{2}) \geq \frac{1}{2}f_{s+1}(n). \tag{4.28}$$

Hence, we can get the desired lemma from cases (1) and (2). □

Proof of Lemma 4.4. Suppose that we have n randomized particles $y_1, y_2, \cdots, y_n$. Each particle is uniformly distributed in $\Lambda_A \cup \Lambda_B$ so that, for $i = 1, 2, \cdots, n$,

$$P(y_i \in \Lambda_A) = P(y_i \in \Lambda_B) = \frac{1}{2}. \tag{4.29}$$

Let

$$m_A = \sum_{i=1}^{n} 1(y_i \in \Lambda_A), \quad m_B = \sum_{i=1}^{k} 1(y_i \in \Lambda_B), \tag{4.30}$$

and $\langle\cdot\rangle$ denote expectation with respect to the distribution of $y_1, y_2, \cdots, y_n$. Then,

$$F_s(0, 0, n) = \frac{\langle f_s(m_A) + f_s(m_B)\rangle}{|\Lambda_A|}. \tag{4.31}$$

To compute $\langle f_s(m_A)\rangle$, we first calculate $\langle m_A^2 - m_A\rangle$ and estimate the difference $\langle m_A^2 - m_A - f_s(m_A)\rangle$. The former is

$$\langle m_A^2 - m_A\rangle = \langle m_A\rangle^2 + (\langle m_A^2\rangle - \langle m_A\rangle^2) - \langle m_A\rangle = \frac{n^2}{4} + \frac{n}{4} - \frac{n}{2} = \frac{n^2}{4} - \frac{n}{4}. \tag{4.32}$$

To estimate the latter, we use the Chernoff bound (See Corollary 4.9 in [15].) for binomial distribution, which becomes, in this case,

$$P(m_A \geq n/2 + \zeta) \leq e^{-\frac{2\zeta^2}{n}}. \tag{4.33}$$

where we used the mean and variance of m_A are $n/2$ and $n/4$. Since

$$t^2 - t - f(x) = \begin{cases} (t - K_s)^2 & \text{if } t > K_s \\ 0 & \text{if } t \leq K_s \end{cases}, \tag{4.34}$$

(4.33) implies

$$\begin{aligned} \langle m_A^2 - m_A - f_s(m_A)\rangle &= \sum_{\zeta=K_s}^{\infty} P(m_A = \zeta)\cdot(\zeta - K_s)^2 \leq \sum_{\zeta=K_s}^{\infty} P(m_A \geq \zeta)\cdot(\zeta - K_s)^2 \\ \leq \quad \sum_{\zeta=K_s}^{\infty} e^{-\frac{2(\zeta-n/2)^2}{n}}(\zeta - K_s)^2 &\leq \int_{K_s}^{\infty} e^{-\frac{2(\zeta-n/2)^2}{n}}(\zeta - K_s)^2 d\zeta \leq Cn^2 e^{-\frac{(K_s-n/2)^2}{n}}. \end{aligned} \tag{4.35}$$

To estimate the last term in the inequality above, we use $n < K_{s+1} + 2\sqrt{K_{s+1}}$ and (4.8) and obtain

$$K_s - \frac{n}{2} \quad \geq \quad K_s - \frac{K_{s+1}}{2} - \sqrt{K_{s+1}} \gg |\log\rho|^{\frac{1}{2}}\sqrt{K_{s+1}}. \tag{4.36}$$

Hence,

$$\frac{(K_s - n/2)^2}{n} \geq \frac{(K_s - n/2)^2}{2^{s+2}\rho\ell_1^3} \gg |\log\rho|, \tag{4.37}$$

which gives

$$n^2 e^{-\frac{(K_s-n/2)^2}{n}} \ll \rho n^2. \tag{4.38}$$

Together with (4.35) we get

$$\langle m_A^2 - m_A - f_s(m_A)\rangle \ll \rho n^2, \tag{4.39}$$

thus,

$$\langle f_s(m_A)\rangle \geq (1 - \rho)(\frac{n^2}{4} - \frac{n}{4}). \tag{4.40}$$

Therefore, from that

$$\langle f_s(m_A) + f_s(m_B)\rangle = 2\langle f_s(m_A)\rangle, \tag{4.41}$$

we obtain, when $n < K_{s+1} + 2\sqrt{K_{s+1}}$,

$$F_s(0,0,n) \geq \frac{1}{2}(1-\rho)\frac{n^2-n}{|\Lambda_A|} = (1-\rho)\frac{n^2-n}{|\Lambda_A \cup \Lambda_B|} \geq (1-\rho)\frac{f_{s+1}(n)}{|\Lambda_A \cup \Lambda_B|}, \tag{4.42}$$

which was to be proved. □

4.4. **Proof of Lemma 4.5.** In this subsection, we consider s to be fixed and let $\ell = \ell(s)$, $F = F_s$ and $n < K_{s+1} + 2\sqrt{K_{s+1}}$. We will see that the ground state energy is equal to the expectation with respect to the uniform distribution of particles up to a small error. And this is guaranteed by the small portion of the kinetic energies $\sum T_j$. The following lemmas show the idea:

Lemma 4.6. *When $n < K_{s+1} + 2\sqrt{K_{s+1}}$, for any fixed $x_2, x_3, \cdots, x_N$ with $x_{n+1}, x_{n+2}, \cdots, x_N$ outside $(\Lambda_A \cup \Lambda_B)$, assume n_A particles among $x_2, \cdots, x_n$ are in Λ_A and n_B particles in $\Lambda_B (n_A + n_B = n-1)$. So, when x_1 is in Λ_A the total energy is $F(n_A+1, n_B, 0)$, otherwise it is $F(n_A, n_B+1, 0)$. Let*

$$T_1 = -W_1^{-1}\nabla_1 W_1^2 \nabla_1 W_1^{-1} \tag{4.43}$$

defined on functions in $L^2(\Lambda_A \cup \Lambda_B)$ with the W_1-Neumann boundary conditions in $\Lambda_A \cup \Lambda_B$. Then, there exists a constant C' such that, we have the following operator inequality:

$$\begin{aligned} &\frac{(4\pi a)^{-1}}{|\log\rho|}T_1 + F(n_A+1, n_B, 0)\cdot 1(x_1 \in \Lambda_A) + F(n_A, n_B+1, 0)\cdot 1(x_1 \in \Lambda_B) \\ \geq\ & F(n_A, n_B, 1) - C'\frac{(n_A - n_B)^2}{\ell^4 |\log\rho|^{-1}} - Cn\rho^{\frac{1}{3}}\ell^{-3}. \end{aligned} \tag{4.44}$$

We note that the LHS of (4.44) depends on x_1, but the RHS of (4.44) does not. Instead, the function F in the right hand side has 1 instead of 0 in its third argument, which means that, when we compute F, we need to consider one imaginary particle whose distribution is uniform in $\Lambda_A \cup \Lambda_B$ as in the definition of F. Thus, we can say that the particle x_1 got randomized.

We note that the right hand side has a term $-C'(n_A - n_B)^2 \ell^{-4}|\log\rho|$. When we apply perturbation theory with x_k, $2 \leq k \leq n$, we also need to take this term into consideration. The following lemma shows an outcome of perturbation theory for general k:

Lemma 4.7. *With the assumption as above, except that n_A particles among $x_{k+1}, \cdots, x_n$ are in Λ_A and n_B particles in Λ_B $(n_A + n_B = n-k)$, we have*

$$\begin{aligned} &\frac{(4\pi a)^{-1}}{|\log\rho|}T_k + \left(F(n_A+1, n_B, k-1) - C'(k-1)\frac{(n_A+1-n_B)^2}{\ell^4|\log\rho|^{-1}}\right)\cdot 1(x_k \in \Lambda_A) \\ &\qquad + \left(F(n_A, n_B+1, k-1) - C'(k-1)\frac{(n_A-n_B-1)^2}{\ell^4|\log\rho|^{-1}}\right)\cdot 1(x_k \in \Lambda_B) \\ \geq\ & F(n_A, n_B, k) - C'k\frac{(n_A-n_B)^2}{\ell^4|\log\rho|^{-1}} - Cn\rho^{\frac{1}{3}}\ell^{-3}. \end{aligned} \tag{4.45}$$

Here, C does not depend on k and T_k is defined as

$$T_k = -W_k^{-1}\nabla_k W_k^2 \nabla_k W_k^{-1} \tag{4.46}$$

We can prove Lemma 4.5 using the above lemma n times.

Proof of Lemma 4.5. Let

$$\mathcal{M}^{(k)}(A) = \sum_{i=k+1}^{n} 1(x_i \in \Lambda_A), \quad \mathcal{M}^{(k)}(B) = \sum_{i=k+1}^{n} 1(x_i \in \Lambda_B). \tag{4.47}$$

Note that $\mathcal{M}(A)=\mathcal{M}^{(0)}(A)$ and $\mathcal{M}(B)=\mathcal{M}^{(0)}(B)$.

From Lemma 4.6, we get

$$\begin{aligned}&\frac{(4\pi a)^{-1}}{|\log\rho|}\int_{(\Lambda_A\cup\Lambda_B)^n}|W_1|^2|\nabla_1\phi_1(\mathbf{x}_n)|^2d\mathbf{x}_n+\int_{(\Lambda_A\cup\Lambda_B)^n}F(\mathcal{M}(A),\mathcal{M}(B),0)|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n \\ \geq\ &\int_{(\Lambda_A\cup\Lambda_B)^n}\big[F(\mathcal{M}^{(1)}(A),\mathcal{M}^{(1)}(B),1)-C'\frac{(\mathcal{M}^{(1)}(A)-\mathcal{M}^{(1)}(B))^2}{\ell^4|\log\rho|^{-1}}-Cn\rho^{\frac13}\ell^{-3}\big]|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n.\end{aligned}\tag{4.48}$$

As a next step, we apply Lemma 4.7 with the right hand side of the above equation. Then we get,

$$\begin{aligned}&\frac{(4\pi a)^{-1}}{|\log\rho|}\int_{(\Lambda_A\cup\Lambda_B)^n}|W_2|^2|\nabla_2\phi_2(\mathbf{x}_n)|^2d\mathbf{x}_n \\ &+\int_{(\Lambda_A\cup\Lambda_B)^n}\big[F(\mathcal{M}^{(1)}(A),\mathcal{M}^{(1)}(B),1)-C'\frac{(\mathcal{M}^{(1)}(A)-\mathcal{M}^{(1)}(B))^2}{\ell^4|\log\rho|^{-1}}\big]|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n \\ \geq\ &\int_{(\Lambda_A\cup\Lambda_B)^n}\big[F(\mathcal{M}^{(2)}(A),\mathcal{M}^{(2)}(B),2)-C'\frac{(\mathcal{M}^{(2)}(A)-\mathcal{M}^{(2)}(B))^2}{\ell^4|\log\rho|^{-1}}-Cn\rho^{\frac13}\ell^{-3}\big]|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n.\end{aligned}\tag{4.49}$$

Thus,

$$\begin{aligned}&\sum_{i=1,2}\frac{(4\pi a)^{-1}}{|\log\rho|}\int_{(\Lambda_A\cup\Lambda_B)^n}|W_i|^2|\nabla_i\phi_i(\mathbf{x}_n)|^2d\mathbf{x}_n+\int_{(\Lambda_A\cup\Lambda_B)^n}F(\mathcal{M}(A),\mathcal{M}(B),0)|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n \\ \geq\ &\int_{(\Lambda_A\cup\Lambda_B)^n}\big[F(\mathcal{M}^{(2)}(A),\mathcal{M}^{(2)}(B),2)-C'\frac{(\mathcal{M}^{(2)}(A)-\mathcal{M}^{(2)}(B))^2}{\ell^4|\log\rho|^{-1}}-Cn\rho^{\frac13}\ell^{-3}\big]|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n.\end{aligned}$$

We keep applying Lemma 4.7. Since $\mathcal{M}^{(k)}(A)=\mathcal{M}^{(k)}(B)=0$, when $k=n$, we get

$$\begin{aligned}&\sum_{i=1}^{n}\frac{(4\pi a)^{-1}}{|\log\rho|}\int_{(\Lambda_A\cup\Lambda_B)^n}|W_i|^2|\nabla_i\phi_i(\mathbf{x}_n)|^2d\mathbf{x}_n+\int_{(\Lambda_A\cup\Lambda_B)^n}F(\mathcal{M}(A),\mathcal{M}(B),0)|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n \\ \geq\ &\int_{(\Lambda_A\cup\Lambda_B)^n}\big[F(\mathcal{M}^{(n)}(A),\mathcal{M}^{(n)}(B),n)-C'\frac{(\mathcal{M}^{(n)}(A)-\mathcal{M}^{(n)}(B))^2}{\ell^4|\log\rho|^{-1}}-Cn^2\rho^{\frac13}\ell^{-3}\big]|\psi(\mathbf{x}_n)|^2d\mathbf{x}_n \\ =\ &\big[F(0,0,n)-Cn^2\rho^{\frac13}\ell^{-3}\big]\|\psi\|_2^2.\end{aligned}\tag{4.50}$$

Since F is convex, from Jensen's inequality,

$$F(0,0,n)\geq F(\frac{n}{2},\frac{n}{2},0)\geq Cn^2\ell^{-3}.\tag{4.51}$$

Thus, we have

$$\begin{aligned}&\frac{(4\pi a)^{-1}}{|\log\rho|}\sum_{j=1}^{n}\int_{(\Lambda_A\cup\Lambda_B)^n}|W_j|^2|\nabla_j\phi_j(\mathbf{x}_n)|^2d\mathbf{x}_n+\int_{(\Lambda_A\cup\Lambda_B)^n}F(\mathcal{M}(A),\mathcal{M}(B),0)|\psi(x_n)|^2d\mathbf{x}_n \\ \geq\ &(1-C\rho^{\frac13})F(0,0,n)\|\psi(x_n)\|_2^2d\mathbf{x}_n,\end{aligned}\tag{4.52}$$

which was to be proved. □

Since Lemma 4.6 is a special case of Lemma 4.7, we only prove Lemma 4.7.

Proof of Lemma 4.7. Let

$$g_k(x_k)=\begin{cases}F(n_A+1,n_B,k-1)-F(n_A,n_B,k-1) & \text{if } x_k\in\Lambda_A\\ F(n_A,n_B+1,k-1)-F(n_A,n_B,k-1) & \text{if } x_k\in\Lambda_B\end{cases}\tag{4.53}$$

and

$$M_k(x_k) = \begin{cases} (n_A+1-n_B)^2-(n_A-n_B)^2 & \text{if } x_k \in \Lambda_A \\ (n_A-n_B-1)^2-(n_A-n_B)^2 & \text{if } x_k \in \Lambda_B \end{cases}. \tag{4.54}$$

Consider

$$\frac{(4\pi a)^{-1}}{|\log\rho|}T_k + g_k - C'(k-1)\frac{M_k}{\ell^4|\log\rho|^{-1}} \tag{4.55}$$

with $(4\pi a|\log\rho|)^{-1}T_k$ as the unperturbed part. We want to use Temple's inequality to get a lower bound, i.e.,

$$\begin{aligned} &\frac{(4\pi a)^{-1}}{|\log\rho|}T_k + g_k - C'(k-1)\frac{M_k}{\ell^4|\log\rho|^{-1}} \\ \geq\ & F(n_A,n_B,k) - F(n_A,n_B,k-1) - \frac{C'}{\ell^4|\log\rho|^{-1}}(n_A-n_B)^2 - Cn\rho^{\frac{1}{3}}\ell^{-3}, \end{aligned} \tag{4.56}$$

which implies Lemma 4.7. So to prove Lemma 4.7, it only remains to prove (4.56).

In order to use Temple's inequality on (4.55), we first check if the perturbation part is non-negative. This is trivial when $k=1$, so we assume that $k \geq 2$. Without loss of generality, assume $x_k \in \Lambda_A$. From definition,

$$\begin{aligned} & g_k(x_k) = F(n_A+1,n_B,k-1) - F(n_A,n_B,k-1) \\ =\ & \langle F(n_A+1+m_A(k-1), n_B+m_B(k-1),0) - F(n_A+m_A(k-1),n_B+m_B(k-1),0)\rangle_{k-1} \\ =\ & \langle \frac{f_s(n_A+1+m_A(k-1))}{|\Lambda_A|} - \frac{f_s(n_A+m_A(k-1))}{|\Lambda_A|}\rangle_{k-1} \end{aligned} \tag{4.57}$$

It can be easily checked that

$$f_s(n_A+1+m_A(k-1)) - f_s(n_A+m_A(k-1)) \geq C(n_A+m_A(k-1)). \tag{4.58}$$

Thus,

$$g_k(x_k) \geq C\ell^{-3}\langle n_A+m_A(k-1)\rangle_{k-1} \geq Ck\ell^{-3}. \tag{4.59}$$

We can prove similarly that $g_k(x_k) \geq Ck\ell^{-3}$ when $x_k \in \Lambda_B$. Since $n \leq C\rho\ell^3 \ll \ell|\log\rho|^{-1}$ and $M_k \leq Cn$, we have

$$g_k \geq Ck\ell^{-3} \gg C\frac{kn}{\ell^4|\log\rho|^{-1}} \geq \frac{kM_k}{\ell^4|\log\rho|^{-1}}, \tag{4.60}$$

which shows that the perturbation part of (4.55) is non-negative, i.e.,

$$g_k - C'k\frac{M_k}{\ell^4|\log\rho|^{-1}} \geq 0. \tag{4.61}$$

We also know the gap of $(4\pi a|\log\rho|)^{-1}T_k$ is much larger than the expectation value of g_k in the ground state W_k, since

$$\ell^{-2}|\log\rho|^{-1} \gg Cn\ell^{-3} \geq \|g_k\|_\infty \geq \langle g_k\rangle_{W_k}. \tag{4.62}$$

Hence, we can apply Temple's inequality on (4.55) and obtain

$$\begin{aligned} &\frac{(4\pi a)^{-1}}{|\log\rho|}T_k + g_k - C'(k-1)\frac{M_k}{\ell^4|\log\rho|^{-1}} \\ \geq\ & \langle g_k\rangle_{W_k} - C'(k-1)\langle\frac{M_k}{\ell^4|\log\rho|^{-1}}\rangle_{W_k} - C\frac{\langle g_k^2\rangle_{W_k} - \langle g_k\rangle_{W_k}^2}{\ell^{-2}|\log\rho|^{-1}} - Ck^2\frac{\langle M_k^2\rangle_{W_k} - \langle M_k\rangle_{W_k}^2}{\ell^8|\log\rho|^{-2}}\Big/\big(\ell^2|\log\rho|^{-1}\big), \end{aligned} \tag{4.63}$$

where $\langle g_k\rangle_{W_k}$ denotes

$$\langle g_k\rangle_{W_k} = \Big(\int_{\Lambda_A\cup\Lambda_B} g_k(x_k)|W_k|^2dx_k\Big)\Big/\Big(\int_{\Lambda_A\cup\Lambda_B}|W_k|^2dx_k\Big) \tag{4.64}$$

and other expectations are defined similarly.

We want to estimate terms in the right hand side of (4.63). In each estimate, we want to compare the expectation $\langle\cdot\rangle_{W_k}$ with an expectation with respect to a uniform distribution. Let $\langle g_k\rangle_{\mathbf{1}_k}$ denote

$$\langle g_k\rangle_{\mathbf{1}_k} = \Big(\int_{\Lambda_A\cup\Lambda_B} g_k(x_k)\cdot 1\, dx_k\Big)\Big/\Big(\int_{\Lambda_A\cup\Lambda_B} 1\, dx_k\Big) \tag{4.65}$$

and other expectations $\langle\cdot\rangle_{\mathbf{1}_k}$ are denoted similarly. It can be easily checked from the definitions that

$$F(n_A, n_B, k-1) + \langle g_k\rangle_{\mathbf{1}_k} = F(n_A, n_B, k). \tag{4.66}$$

(1) Using Lemma 5.3, we get

$$|\langle g_k\rangle_{W_k} - \langle g_k\rangle_{\mathbf{1}_k}| \le Cn\rho\ell_0^2\ell^{-3}. \tag{4.67}$$

Together with (4.66), we have

$$\langle g_k\rangle_{W_k} \ge F(n_A, n_B, k) - F(n_A, n_B, k-1) - Cn\rho^{\frac{1}{3}}\ell^{-3}. \tag{4.68}$$

(2) Using Lemma 5.3, we get

$$k\Big|\langle\frac{M_k}{\ell^4|\log\rho|^{-1}}\rangle_{W_k} - \langle\frac{M_k}{\ell^4|\log\rho|^{-1}}\rangle_{\mathbf{1}_k}\Big| \le Ck\frac{n\rho\ell_0^2}{\ell^4|\log\rho|^{-1}} \ll n\rho\ell_0^2\ell^{-3}. \tag{4.69}$$

On the other hand, we simply follow the definitions to see

$$\langle M_k\rangle_{\mathbf{1}_k} = 1. \tag{4.70}$$

Hence, we get

$$C'(k-1)\langle\frac{M_k}{\ell^4|\log\rho|^{-1}}\rangle_{W_k} \le Cn\rho^{\frac{1}{3}}\ell^{-3}. \tag{4.71}$$

(3) Using Lemma 5.3, we get

$$|\langle g_k^2\rangle_{W_k} - \langle g_k^2\rangle_{\mathbf{1}_k}| \le Cn^2\rho\ell_0^2\ell^{-6}, \tag{4.72}$$

$$|\langle g_k\rangle_{W_k}^2 - \langle g_k\rangle_{\mathbf{1}_k}^2| \le Cn^2\rho\ell_0^2\ell^{-6}. \tag{4.73}$$

Thus, since $n \le C\rho\ell^3 \ll \ell|\log\rho|^{-1}$,

$$\frac{\langle g_k^2\rangle_{W_k} - \langle g_k\rangle_{W_k}^2}{\ell^{-2}|\log\rho|^{-1}} \le \frac{\langle g_k^2\rangle_{\mathbf{1}_k} - \langle g_k\rangle_{\mathbf{1}_k}^2}{\ell^{-2}|\log\rho|^{-1}} + Cn\rho\ell_0^2\ell^{-3}. \tag{4.74}$$

To estimate $\langle g_k^2\rangle_{\mathbf{1}_k} - \langle g_k\rangle_{\mathbf{1}_k}^2$, we again follow the definition and get

$$\langle g_k^2\rangle_{\mathbf{1}_k} - \langle g_k\rangle_{\mathbf{1}_k}^2 = \frac{1}{4}\big[F(n_A+1, n_B, k-1) - F(n_A, n_B+1, k-1)\big]^2. \tag{4.75}$$

Suppose that we have $(k-1)$ particles randomized, namely $y_1, y_2, \cdots, y_{k-1}$ as in the definition of F, (4.17)-(4.20). Then,

$$\begin{aligned} &\big[F(n_A+1, n_B, k-1) - F(n_A, n_B+1, k-1)\big]^2 \qquad (4.76)\\ = \;& \big[\langle F(n_A+1+m_A(k-1), n_B+m_B(k-1), 0) - F(n_A+m_A(k-1), n_B+1+m_B(k-1), 0)\rangle_{k-1}\big]^2\\ \le \;& \big\langle\big[F(n_A+1+m_A(k-1), n_B+m_B(k-1), 0) - F(n_A+m_A(k-1), n_B+1+m_B(k-1), 0)\big]^2\big\rangle_{k-1} \end{aligned}$$

It can be easily checked that, for any non-negative integers m_1 and m_2,

$$[F(m_1+1, m_2, 0) - F(m_1, m_2+1, 0)]^2 \le \frac{4(m_1-m_2)^2}{|\Lambda_A|}. \tag{4.77}$$

Thus,

$$\begin{aligned} &\big[F(n_A+1, n_B, k-1) - F(n_A, n_B+1, k-1)\big]^2\\ \le \;& 4|\Lambda_A|^{-2}\big\langle\big[(n_A+m_A(k-1)) - (n_B+m_B(k-1))^2\big]\big\rangle_{k-1}. \qquad (4.78) \end{aligned}$$

From the definition,

$$\left\langle \left[(n_A + m_A(k-1)) - (n_B + m_B(k-1))^2 \right] \right\rangle_{k-1} = (n_A - n_B)^2 + (k-1). \tag{4.79}$$

Thus, from (4.74), (4.75), (4.78), and (4.79), we obtain

$$\frac{\langle g_k^2 \rangle_{W_k} - \langle g_k \rangle_{W_k}^2}{\ell^{-2} |\log \rho|^{-1}} \le |\Lambda_A|^{-2} \frac{(n_A - n_B)^2 + k}{\ell^{-2} |\log \rho|^{-1}} \le C \frac{(n_A - n_B)^2}{\ell^4 |\log \rho|^{-1}} + C n \rho^{\frac{1}{3}} \ell^{-3}. \tag{4.80}$$

(4) Using Lemma 5.3, from that $n \le C\rho\ell^3 \ll \ell |\log \rho|^{-1}$ we get

$$k^2 \left| \frac{\langle M_k^2 \rangle_{W_k} - \langle M_k \rangle_{W_k^2}}{\ell^6 |\log \rho|^{-3}} - \frac{\langle M_k^2 \rangle_{\mathbf{1}_k} - \langle M_k \rangle_{\mathbf{1}_k}^2}{\ell^6 |\log \rho|^{-3}} \right| \le C k^2 \frac{n^2 \rho \ell_0^2}{\ell^6 |\log \rho|^{-3}} \ll n \rho \ell_0^2 \ell^{-3}. \tag{4.81}$$

From the definition of M_k, it can be easily checked that

$$\langle M_k^2 \rangle_{\mathbf{1}_k} - \langle M_k \rangle_{\mathbf{1}_k}^2 = 4(n_A - n_B)^2. \tag{4.82}$$

Thus,

$$k^2 \frac{\langle M_k^2 \rangle_{\mathbf{1}_k} - \langle M_k \rangle_{\mathbf{1}_k}^2}{\ell^6 |\log \rho|^{-3}} = k^2 \frac{4(n_A - n_B)^2}{\ell^6 |\log \rho|^{-3}} \le \frac{C n^4}{\ell^6 |\log \rho|^{-3}} \ll n \rho^{\frac{1}{3}} \ell^{-3}. \tag{4.83}$$

Here, the last inequality follows from

$$n^3 \ell^{-3} |\log \rho|^3 \le C \rho^3 \ell_2^6 |\log \rho|^3 \ll \rho^{\frac{1}{3}}. \tag{4.84}$$

Hence,

$$k^2 \frac{\langle M_k^2 \rangle_{W_k} - \langle M_k \rangle_{W_k}^2}{\ell^6 |\log \rho|^{-3}} \le n \rho^{\frac{1}{3}} \ell^{-3}. \tag{4.85}$$

Applying estimates (4.68), (4.71), (4.80), and (4.85) to (4.63), we obtain (4.56). This proves the desired lemma. □

5. Properties of two body problem and the approximation to the ground state

Lemma 5.1. *Let e_0 and ϕ be the lowest Neumann eigenvalue and eigenfunction on the ball of radius κ, i.e.,*

$$(-\Delta + \frac{1}{2} V)\phi = e_0 \phi \tag{5.1}$$

with the boundary condition

$$\phi(x) = 1 \text{ and } \partial\phi(x) = 0 \text{ if } |x| = \kappa. \tag{5.2}$$

Suppose that $V(x) = 0$ when $|x| > R_0$. Then, if $\kappa \gg R_0$, there exist constants c_0 and c_1 such that

$$c_0 \le \phi(x) \le 1, \quad 1 - \phi(x) \le \frac{c_1}{|x|}. \tag{5.3}$$

Moreover, we have

$$\frac{3a}{\kappa^3} \le e_0 \le \frac{3a}{\kappa^3}(1 + \frac{C}{\kappa}). \tag{5.4}$$

Proof. For a proof of the first part of the lemma,

$$c_0 \le \phi(x) \le 1, \quad 1 - \phi(x) \le \frac{c_1}{|x|}, \tag{5.5}$$

see Lemma A.1 in [3]. To prove the second part, we extend the definition of ϕ so that $\phi(x) = 1$ if $|x| > \kappa$. It is well known that

$$\int_{\mathbb{R}^3} \phi(x) \left(-\Delta + \frac{1}{2} V(x) \right) \phi(x) dx \ge 4\pi a. \tag{5.6}$$

Thus, $\phi(x) \le 1$ implies that

$$e_0 \ge 4\pi a \big(\int_{|x|\le\kappa} |\phi(x)|^2 dx \big)^{-1} \ge \frac{3a}{\kappa^3}. \tag{5.7}$$

Upper bound for e_0 is also proved in Lemma A.1 in [3]. □

Lemma 5.2. *Define W_j as in (2.14) and $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$. Suppose that x_j is in Λ_ℓ, a box of side length ℓ and $x_1, \cdots, \widehat{x_j}, \cdots, x_N$ are fixed.*

(1) There exists a constant $c_0 < 1$ such that

$$W_j(\boldsymbol{x}) \ge 1 - c_0. \tag{5.8}$$

(2) Let $\ell \gg \ell_0$. If Λ_ℓ contains n particles including x_j, then

$$\ell^3 (1 - Cn\frac{\ell_0^2}{\ell^3} - C\ell^{-1}) \le \int_{\Lambda_\ell} |W_j|^2 dx_j \le \ell^3 \tag{5.9}$$

Proof. ¿From Lemma 5.1, we can see that there exists a constant c_0 that does not depend on κ such that $\tau(\kappa, x_i - x_j) \le c_0$ whenever $\ell_{-1} \le \kappa \le \ell_0$. It is clear that

$$F_{ij}(\mathbf{x})\tau(\ell_0, x_i - x_j) > 0 \text{ or } (1 - F_{ij}(\mathbf{x}))G_{ij}(\mathbf{x})\tau(t_{ij}, x_i - x_j) > 0 \tag{5.10}$$

implies that x_i is the nearest particle of x_j. Thus, there exists an index k such that

$$W_j(\mathbf{x}) = 1 - [F_{kj}(\mathbf{x})\tau(\ell_0, x_k - x_j) + (1 - F_{kj}(\mathbf{x}))G_{kj}(\mathbf{x})\tau(t_{kj}, x_k - x_j)], \tag{5.11}$$

hence, we obtain the first part of the lemma.

To prove the second part of the lemma, we note that $W_j = 1$ unless there exists a particle x_k such that $|x_k - x_j| < \min\{\ell_0, t_{kj}\}$. Assume that there exists such k. Consider first a case where $x_k \in \Lambda_\ell$. When $t_{kj} > \ell_0$,

$$\int_{|x_k - x_j| < \ell_0} |W_j|^2 dx_j \ge \int_{|x_k - x_j| < \ell_0} \big(1 - \frac{C}{|x_k - x_j|}\big)^2 dx_j \ge \int_{|x_k - x_j| < \ell_0} dx_j - C\ell_0^2, \tag{5.12}$$

and, when $t_{kj} \le \ell_0$,

$$\int_{|x_k - x_j| < t_{ij}} |W_j|^2 dx_j \ge \int_{|x_k - x_j| < t_{ij}} \big(1 - \frac{C}{|x_k - x_j|}\big)^2 dx_j \ge \int_{|x_k - x_j| < t_{ij}} dx_j - Ct_{ij}^2 \ge \frac{4}{3}\pi t_{ij}^3 - C\ell_0^2. \tag{5.13}$$

Now consider the other case where $x_k \notin \Lambda_\ell$. Note that $d(x_k, \Lambda_\ell) \le \ell_0$ in this case, where $d(x_k, \Lambda_\ell)$ denotes the distance from x_k to the box Λ_ℓ. When $t_{kj} > \ell_0$, let B_k be a ball of radius ℓ_0 centered at x_k. $B_k \cap \partial\Lambda_\ell$, the intersection of B_k and the boundary of Λ_ℓ, is a disk, and we let r_k be the radius of the disk. Since

$$\int_{B_k \cap \Lambda_\ell} \frac{1}{|x_j - x_k|} dx_j \le \int_{\sqrt{\ell_0^2 - r_k^2} \le |x_j - x_k| \le \ell_0} \frac{1}{|x_j - x_k|} dx_j = 2\pi r_k^2, \tag{5.14}$$

we can see that

$$\int_{B_k \cap \Lambda_\ell} |W_j|^2 dx_j \ge \int_{B_k \cap \Lambda_\ell} \big(1 - \frac{C}{|x_k - x_j|}\big)^2 dx_j \ge \int_{B_k \cap \Lambda_\ell} dx_j - Cr_k^2. \tag{5.15}$$

We can get the same estimate when $t_{kj} \le \ell_0$, in this case, B_k becomes a ball of radius t_{kj} centered at x_k. Note that all these B_k's are disjoint. If we sum over all such k, we get

$$\sum_{k: x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \le \ell_0} \int_{B_k \cap \Lambda_\ell} |W_j|^2 dx_j \ge \sum_{k: x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \le \ell_0} \int_{B_k \cap \Lambda_\ell} dx_j - C \sum_{k: x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \le \ell_0} r_k^2. \tag{5.16}$$

Since $B_k \cap \partial\Lambda_\ell$ lies on $\partial\Lambda_\ell$, $\sum r_k^2$ cannot exceed the area of $\partial\Lambda_\ell$, thus,

$$\sum_{k: x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \le \ell_0} \int_{B_k \cap \Lambda_\ell} |W_j|^2 dx_j \ge \sum_{k: x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \le \ell_0} \int_{B_k \cap \Lambda_\ell} dx_j - C\ell^2. \tag{5.17}$$

Altogether, we get

$$\int_{\Lambda_\ell} |W_j|^2 dx_j \geq \int_{\Lambda_\ell} dx_j - Cn\ell_0^2 - C\ell^2 = \ell^3(1 - Cn\frac{\ell_0^2}{\ell^3} - C\ell^{-1}). \tag{5.18}$$

Finally, $W_j \leq 1$ implies that

$$\int_{\Lambda_\ell} |W_j|^2 dx_j \leq \ell^3. \tag{5.19}$$

This proves the second part of the lemma. □

Lemma 5.3. *Let Λ_A and Λ_B be as in Section 4. Suppose $x_1, x_2, \cdots, x_n \in \Lambda_A \cup \Lambda_B$ and $n = O(\rho|\Lambda_A|)$. Let $1 \leq j \leq n$ and let W_j be defined in (2.14). Fix $x_1, x_2, \cdots, \widehat{x_j}, \cdots, x_N$ with $x_{n+1}, x_{n+2}, \cdots, x_N$ outside $(\Lambda_A \cup \Lambda_B)$. For a given function $h(x_j)$ with*

$$h(x_j) = \begin{cases} h_A & \text{if } x_j \in \Lambda_A \\ h_B & \text{if } x_j \in \Lambda_B \end{cases}, \tag{5.20}$$

let

$$\langle h \rangle_{W_j} = \int_{\Lambda_A \cup \Lambda_B} h(x_j)|W_j|^2 dx_j / \int_{\Lambda_A \cup \Lambda_B} |W_j|^2 dx_j, \tag{5.21}$$

$$\langle h \rangle_{\mathbf{1}_k} = \int_{\Lambda_A \cup \Lambda_B} h(x_j) \cdot 1 \, dx_j / \int_{\Lambda_A \cup \Lambda_B} 1 \, dx_j. \tag{5.22}$$

Then,

$$|\langle h \rangle_{W_j} - \langle h \rangle_{\mathbf{1}_k}| \leq C\rho\ell_0^2 \max\{h_A, h_B\}. \tag{5.23}$$

Proof. Let

$$p_A = (\int_{\Lambda_A} |W_j|^2 dx_j)/(\int_{\Lambda_A \cup \Lambda_B} |W_j|^2 dx_j), \tag{5.24}$$

$$p_B = (\int_{\Lambda_B} |W_j|^2 dx_j)/(\int_{\Lambda_A \cup \Lambda_B} |W_j|^2 dx_j). \tag{5.25}$$

Then,

$$\langle h \rangle_{W_j} = p_A h_A + p_B h_B. \tag{5.26}$$

¿From Lemma 5.2,

$$p_A = \frac{1}{2} + O(\frac{n\ell_0^2}{\ell^3}) + O(\ell^{-1}) = \frac{1}{2} + O(\rho\ell_0^2) + O(\ell^{-1}), \tag{5.27}$$

and the same estimate holds for p_B. Since

$$\langle h \rangle_{\mathbf{1}_k} = \frac{1}{2}h_A + \frac{1}{2}h_B, \tag{5.28}$$

and $\ell^{-1} \ll \rho\ell_0^2$, we get

$$|\langle h \rangle_{W_j} - \langle h \rangle_{\mathbf{1}_k}| \leq C\rho\ell_0^2 h_A + C\rho\ell_0^2 h_B \leq C\rho\ell_0^2 \max\{h_A, h_B\}, \tag{5.29}$$

which was to be proved. □

Acknowledgment

We are grateful to H.-T. Yau for helpful discussions.

References

[1] Davis, K. B., Mewes, M.-O., Andrews, M. R., van Druten, N. J., Durfee, D. S., Kurn, D. M., and Ketterle, W.: *Bose-Einstein Condensation in a Gas of Sodium Atoms*, Phys. Rev. Lett. **75**, 3969-3973 (1995).

[2] Dyson, F. J.: *Ground-State of a Hard-Sphere Gas*, Phys. Rev. **106**, 20-26 (1957).

[3] Erdos, L., Schlein, B., Yau, H.-T.: *Derivation of the Gross-Pitaevskii Hierarchy for the Dynamics of Bose-Einstein Condensate*, Commmun. Pure. Appl. Math. **59**, 1659-1741 (2005).

[4] Lee, J. O.: *Ground State Energy of Dilute Bose Gas in Small Negative Potential Case*, J. Stat. Phys. **134**, 1-18 (2009).

[5] Lee, T. D., Huang, K., Yang, C. N.: *Eigenvalues and Eigenfunctions of a Bose System of Hard Spheres and Its Low-Temperature Properties*, Phys. Rev. **106**, 1135-1145 (1957).

[6] Lee, T. D., Yang, C. N.: *Many-Body Problem in Quantum Mechanics and Quantum Statistical Mechanics*, Phys. Rev. **105**, 1119 (1957).

[7] Lee, T. D., Yang, C. N.: *Many-Body Problem in Quantum Statistical Mechanics. III. Zero-Temperature Limit for Dilute Hard Spheres*, Phys. Rev. **117**, 12-21 (1960).

[8] Lieb, E. H.: *Simplified Approach to the Ground State Energy of an Imperfect Bose Gas*, Phys. Rev. **130**, 2518-2528 (1963).

[9] Lieb, E. H., Seiringer, R., Yngvason, J.: *One-Dimensional Behavior of Dilute, Trapped Bose Gases*, Commun. Math. Phys. **244**, 347-393 (2004).

[10] Lieb, E. H., Solovej, J. P.: *Ground State Energy of the One-Component Charged Bose Gas*, Commun. Math. Phys. **217**, 127-163 (2001). Errata **225**, 219-221 (2002).

[11] Lieb, E. H., Solovej, J. P.: *Ground State Energy of the Two-Component Charged Bose Gas*, Commun. Math. Phys. **252**, 485-534 (2004).

[12] Lieb, E. H., Yngvason, J.: *Ground State Energy of the Low Density Bose Gas*, Phys. Rev. Lett. **80**, 2504-2507 (1998).

[13] Lieb, E. H., Yngvason, J.: *Ground State Energy of a Dilute Two-Dimensional Bose Gas*, J. Stat. Phys. **103**, 509-526 (2001).

[14] Lieb, E. H., Seiringer, R., Solovej, J. P., Yngvason, J.: *The Ground State of the Bose Gas* in *Current Developments in Mathematics, 2001*, International Press, Cambridge, 2002, pp. 131-178, arXiv:math-ph/0204027 v2 (2003).

[15] Mitzenmacher, M., Upfal, E.: *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, Cambridge University Press, New York (2005).

[16] Solovej, J. P.: *Upper Bounds to the Ground State Energies of the One- and Two-Component Charged Bose Gases*, Commun. Math. Phys. **266**, 797-818 (2006).

[17] Schnee, K., Yngvason, J.: *Bosons in Disc-Shaped Traps: From 3D to 2D*, Commun. Math. Phys. **269**, 659-691 (2007).

[18] Temple, G.: *The Theory of Rayleigh's Principle as Applied to Continuous Systems*, Proc. Roy. Soc. London A **119**, 276-293 (1928).

[19] Yau, H.-T., Yin, J.: *The Second Order Upper Bound for the Ground Energy of a Bose Gas*, arXiv:math-ph/0903.5347v1 (2008).

[20] Yin, J.: *The Ground State Energy of Dilute Bose Gas in Potentials with Positive Scattering Length*, arxiv:math-ph/0808.4066v1 (2008).

(Ji Oon Lee) Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
E-mail address: `jioon@math.harvard.edu`

(Jun Yin) Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
E-mail address: `jyin@math.harvard.edu`