A Lower Bound on the Ground State Energy of Dilute Bose Gas

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Lee, Ji, and Jun Yin. 2010. A lower bound on the ground state energy of dilute bose gas. Journal of Mathematical Physics 51(5): 053302.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1063/1.3376639</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:8001961">http://nrs.harvard.edu/urn-3:HUL.InstRepos:8001961</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP</a></td>
</tr>
</tbody>
</table>
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 \( \Lambda \) 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
\[
\frac{E(N, \Lambda)}{N} \geq 4\pi a \rho \left[ 1 - O(\rho^{1/3} |\log \rho|^{3/2}) \right],
\]
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 \[6, 2, 8, 9\], and since the first experimental observation of Bose-Einstein Condensation in 1995 \[1\], interest in low temperature Bose gases are renewed \[13, 14, 11, 12, 10, 17, 18\]. 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 \( \rho \) 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 \[13\]. It is also proved that the leading term is the same in some cases where the interaction potential is not purely non-negative \[5, 21\].

Lee and Yang \[7, 8\] first predicted the second order correction to this leading term, which is given by
\[
4\pi a \rho \left[ 1 + \frac{128}{15\sqrt{\pi}} \left( \rho a^3 \right)^{3/2} + \cdots \right].
\] (1.1)

This calculation used pseudopotential method and binary collision expansion method \[6\]. Another derivation was later given by Lieb \[9\] using a self-consistent closure assumption for the hierarchy of correlation functions.

The matching upper bound of this second order correction is recently obtained in Yau and Yin’s work \[20\]. The matching lower bound, however, has not been proved yet, and in fact, the best result so far is \( 4\pi a \rho \left( 1 - C \rho^{1/17} \right) \), which was first derived by Lieb and Yngvason \[13\]. (See also \[15\] for details.) It also should be mentioned that this second order correction is proved for the high density and weak coupling regime recently by Giuliani and Seiringer \[4\].

Lieb and Yngvason \[13\] 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 \[19\]. 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 \( \varphi \) satisfies

\[
\left( -\Delta + \frac{1}{2} V(x) \right) \varphi(x) = 0.
\]

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

\[
a := \lim_{|x| \to \infty} |x| (1 - \varphi(x))
\]

as in [13].

Hamiltonian of the Bose gas system of \( N \) particles in a torus \( \Lambda \) of side length \( L \) is given by

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

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

\[
x_n = (x_1, x_2, \cdots, x_n), \quad x = x_N.
\]

Throughout the paper, \( C \) denotes a constant independent of \( \rho \).

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

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

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

\[
E(n, \Lambda) := \inf_{\|\psi\|_2 = 1} \left[ \sum_{j=1}^{n} \int_{\Lambda^N} |\nabla_j \psi(x)|^2 dx + \sum_{i<j}^{n} \int_{\Lambda^N} V(x_i - x_j) |\psi(x)|^2 dx \right].
\]

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

\[
\lim_{N, L \to \infty \atop N/L^3 = \rho \to 0} \frac{E(N, \Lambda)}{N} \geq 4\pi a \rho (1 - C_0 \rho^{\frac{1}{2}} |\log \rho|^3)
\]

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.,

\[
(-\Delta + \frac{1}{2}V(x))(1 - \tau(\kappa, x)) = e_0(\kappa)(1 - \tau(\kappa, x))
\]

with the boundary conditions

\[
\tau(\kappa, x) = 0, \quad \partial\tau(\kappa, x) = 0, \quad \text{if } |x| = \kappa.
\]

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

\[
\frac{3a}{\kappa^3} \leq e_0 \leq \frac{3a}{\kappa^3}(1 + \frac{C}{\kappa}).
\]

Using \(\tau\), we define a function \(W_j\) that will be used for our proof, which will approximate the ground state of \(-\Delta_j + \sum_{i\neq j} V(x_i - x_j)\) when \(x\) 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.,

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

We introduce a particle triple cutoff function \(F_{ij}(x)\) at a length scale \(\ell_0\) defined by

\[
F_{ij}(x) = \begin{cases} 
1 & \text{if } t_{ij} > \ell_0, \\
0 & \text{otherwise}.
\end{cases}
\]

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

\[
G_{ij}(x) = \begin{cases} 
1 & \text{if } t_{ij} > \ell_-, \\
0 & \text{otherwise}.
\end{cases}
\]

Here, \(\ell_0\) and \(\ell_-\) are fixed and satisfy \(\ell_0 \gg \ell_- \gg R_0\). 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 \(\tau\) so that \(\tau(\kappa, x) = 0\) if \(|x| > \kappa\). We define \(W_j\) by

\[
W_j(x) = 1 - \sum_{i\neq j} [F_{ij}(x)\tau(t_0, x_i - x_j) + (1 - F_{ij}(x))G_{ij}(x)\tau(t_{ij}, x_i - x_j)].
\]

Our interpretation of this wave function \(W_j\) is as follows: When \(x_i\) is the nearest particle of \(x_j\), \(W_j\) becomes \(1 - \tau(t_0, x_i - x_j)\) 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 \(\ell_0\) or \(t_{ij}\), provided that \(t_{ij} \geq \ell_-\). \(x_j\) 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 \(\tau\).

We will use length scales from now on such that

\[
\ell_- \sim \rho^{-\frac{3}{4}} |\log \rho|^{\frac{3}{4}}, \quad \ell_0 \sim \rho^{-\frac{1}{4}} |\log \rho|^{-\frac{1}{4}}, \quad \ell_1 \sim \rho^{-\frac{1}{4}} |\log \rho|^{\frac{1}{4} + \eta}, \quad \ell_2 \sim \rho^{-\frac{1}{4}} |\log \rho|^{-\frac{3}{4}}, \quad \epsilon = \rho^{\frac{1}{4}} |\log \rho|^3.
\]

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

\[
L = h'\ell_2, \quad \ell_2 = 2^h \ell_1,
\]

where \(h, h'\) are integers.

We will prove in Section 5 that there exists a constant \(c_0 < 1\) such that

\[
W_j(x) \geq 1 - c_0 > 0.
\]
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 $j = 1, 2, \cdots, N$. Here, $\Phi_j$ 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:

**Lemma 2.3.** Let

$$q(\kappa, x) := e_0(\kappa)1(\|x\| \leq \kappa)$$

and

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

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

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

where $\Phi_j(x) := W_j^{-1}\Psi(x)$ 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 $|W_j|^2|\nabla_j \Phi_j(x)|^2$, 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 $U \subset \Lambda$ be a box.

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

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

$$\psi_j(x_j) := \psi(x_n)$$

satisfies $W_j$-Neumann boundary condition in $U$ as defined above.

When we prove lemmas containing $|W_j|^2|\nabla_j \Phi_j(x)|^2$ such as Lemma 2.6 by using $T_j$, we will only consider cases where $\Phi_j(x)$ satisfies $W_j$-Neumann boundary condition, which can be justified as follows: For a given $\psi(x_n) \in L^2(U^n)$, we can find $\tilde{\psi}(x_n) \in L^2(U^n)$ such that $\tilde{\psi}$ satisfies $W_j$-Neumann boundary condition in $U$ for $j = 1, 2, \cdots, n$, and the difference between the kinetic energies of $\psi$ and $\tilde{\psi}$ is negligible. To see this, we first change $\psi$ into $\tilde{\psi}$ so that $\nabla_j(W_j^{-1}\psi(x_j)) = 0$ at the boundary and $\tilde{\psi}$ differs from $\psi$ only near the boundary. If this change was small enough and made sufficiently near the boundary, then, even though $\nabla_j \tilde{\psi}$ might have been changed a lot, the difference between $\nabla_j \psi$ and $\nabla_j \tilde{\psi}$ could be arbitrarily small after integration. This results in the negligible difference between the kinetic energies of $\psi$ and $\tilde{\psi}$.

(2) To use perturbation theory, we divide the torus $\Lambda$ into small cubic cells. We need to replace the soft potential $q_{ij}$ by another soft potential $\tilde{q}_{ij}$ 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 $u$. More precisely, we will consider the following:
Let \( u \in [-\ell_1/2, \ell_1/2]^3 \) and 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 \( \ell_1 \) centered at \((u + \lambda)\). Here, \( u \) corresponds to the origin of the grid that divides \( \Lambda \) into small boxes \( \Lambda_{u\lambda} \). 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 \( \Lambda \) into small cubic cells of side length \( \ell_1 \sim \rho^{-\frac{3}{4}} \log \rho \eta^{-1} n^{\frac{1}{4}} \), \( 0 < \eta < 1/15 \) such that

\[
\Lambda = \bigcup_{\lambda} \Lambda_{u\lambda}, \quad \Lambda_{u\lambda} \cap \Lambda_{u\lambda'} = \emptyset \text{ if } \lambda_i \neq \lambda_j,
\]

where \( \Lambda_{u\lambda} \) is a cubic cell of side length \( \ell_1 \) and \( \lambda \) 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 \( \Lambda \). Define

\[
\tilde{q}_{ij} := \begin{cases} q_{ij} & \text{if } d(x_i, \partial \Lambda_{u\lambda}) > 2\ell_0 \text{ for any } \lambda \\ 0 & \text{otherwise} \end{cases}
\]

Then, for any \( \Psi(x) = \Psi(x_1, x_2, \ldots, x_N) \in L^2(\Lambda^N) \),

\[
\sum_{i \neq j}^{N} \int_{\Lambda^N} q_{ij}(x)|\Psi(x)|^2 dx \geq \inf_u \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \sum_{i \neq j}^{N} \int_{\Lambda^N} \tilde{q}_{ij}(x)|\Psi(x)|^2 dx.
\]

(3) In each small cubic cell of side length \( \ell_1 \), 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 \( x_1, x_2, \ldots, x_N \in \Lambda \). Let \( B \subset \Lambda \) be a box of side length \( \ell_1 \). Define \( E(n, B) \) as the ground state energy of \( n \) particle system in \( B \) with interaction \( V \), i.e.,

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

with the Neumann boundary conditions. Let

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

Let \( \epsilon = \rho \frac{3}{4} \log \rho \eta^3 \) as in (2.14). Fix \( x_{n+1}, x_{n+2}, \ldots, x_N \) outside \( B \) and let \( x_n = (x_1, x_2, \ldots, x_n) \). Then, for any \( \psi(x_n) \in L^2(B^n) \) with \( \phi_j := W_j^{-1} \psi \) for \( j = 1, 2, \ldots, n \), we have

\[
\epsilon E(n, B)||\psi||_2^2 + \frac{1}{2} \sum_{j=1}^{n} \int_{B^n} |W_j| |\nabla_j \phi_j(x_n)|^2 dx_n + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \sum_{i \neq j}^{n} \int_{B^n} \tilde{q}_{ij} |\psi(x_n)|^2 dx_n \geq f(n)(1 - C\epsilon)\frac{4\pi a}{\ell_1^3} ||\psi||_2^2.
\]

(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 \( (1 - \rho^{-1} \ell_1^{-3}) \), which is the ratio between \( n(n-1) \) and \( n^2 \). This error factor becomes quite problematic because \( \rho^{-1} \ell_1^{-3} \sim |\log \rho|^{-1-3\eta} \gg \epsilon \). To handle this problem, we merge \( 2^{3h} \) adjacent small cubic cells of side length \( \ell_1 \) into a larger cubic box of side length \( \ell_2 \). (Recall that we chose \( \ell_2 = 2^b \ell_1 \).) 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 \( \ell_2 \), we can obtain the following lemma:
Lemma 2.7. Let \( x_1, x_2, \ldots, x_N \in \Lambda \). Let \( \Lambda_{t_2} \subset \Lambda \) be a cubic box of side length \( t_2 \sim \rho^{-\frac{1}{3}} |\log \rho|^{-\frac{1}{3}} \). Divide \( \Lambda_{t_2} \) into \( 2^{3h} \) smaller cubic cells of side length \( t_1 \) such that \( t_1 = 2^{-h} t_2 \). Call those small cubic cells \( B_1, B_2, \ldots, B_{2^{3h}} \). Define \( f \) as in (2.20) and let

\[
\tilde{f}(t) := \begin{cases} t(t-1) & \text{if } t \leq \rho t_2^3 \\ (2\rho t_2^3 - 1)t - (\rho t_2^3)^2 & \text{if } t > \rho t_2^3. \end{cases}
\]  

(2.28)

Fix \( x_{n+1}, x_{n+2}, \ldots, x_N \) outside \( \Lambda_{t_2} \) and let \( x_n = (x_1, x_2, \ldots, x_n) \). Define \( N(B_k) \) as

\[
N(B_k) := \sum_{j=1}^n 1(x_j \in B_k).
\]  

(2.29)

Then, for any \( \psi(x_n) \in L^2(\Lambda_{t_2}^n) \) with \( \phi_j = W_j^{-1} \psi \) for \( j = 1, 2, \ldots, n \),

\[
\frac{1}{2} - \epsilon \sum_{j=1}^n \int_{\Lambda_{t_2}^n} \left| W_j \right|^2 \left| \nabla_j \phi_j(x_n) \right|^2 dx_n + \frac{4\pi a}{\ell_1^3} \sum_{k=1}^{2^{3h}} \int_{\Lambda_{t_2}^n} f(|\Psi|) \left| \psi(x_n) \right|^2 dx_n 
\geq \tilde{f}(n)(1 - C \rho^\frac{1}{3} |\log \rho|) \frac{4\pi a}{\ell_3^3} \| \psi \|_2^2.
\]  

(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 \( \rho \) is sufficiently small. Lemma 2.3 shows that, for any \( \Psi \in L^2(\Lambda^N) \),

\[
\langle \Psi, H_N \Psi \rangle \geq \epsilon \langle \Psi, H_N \Psi \rangle + (1 - \epsilon) \left( \sum_{j=1}^N \int_{\Lambda_N} |W_j(x_j)|^2 |\nabla_j \Phi_j(x_j)|^2 dx_j + \sum_{j \neq k}^N \int_{\Lambda_N} q_{jk} |\Psi(x_j)|^2 dx_j \right),
\]  

(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 \( \Lambda \) into small cubic cells \( \Lambda_{u, \lambda} \) of side length \( \ell_1 \), where \( u \) is the origin of division and \( \lambda \in G \) is an index for those cells. From Lemma 2.6, we get

\[
\langle \Psi, H_N \Psi \rangle 
\geq \epsilon \langle \Psi, H_N \Psi \rangle + (1 - \epsilon) \sum_{j=1}^N \int_{\Lambda_N} |W_j(x_j)|^2 |\nabla_j \Phi_j(x_j)|^2 dx_j
\]  

(2.32)

\[
+ (1 - \epsilon) \inf_u \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \sum_{i \neq j}^N \int_{\Lambda_{N-2}} dx_1 dx_2 \cdots dx_i \cdots dx_{i-1} \cdots dx_j \cdots dx_N \left( \sum_{\lambda} \int_{\Lambda_{t_2}^n} \tilde{q}_{ij} |\Psi(x_j)|^2 dx_i dx_j \right).
\]

The next step is to divide the kinetic energy into the small cubic cells. Let \( P \) be the set of functions that assign particles to the small cubic cells, i.e.,

\[
P = \{ p \mid p : \{1, 2, \ldots, N\} \to G \}.
\]  

(2.33)
Let \( p^{-1}(\lambda) = \{ i : p(i) = \lambda, 1 \leq i \leq N \} \). Then,

\[
\sum_{j=1}^{N} \int_{\Lambda^{N}} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 dx = \sum_{j=1}^{N} \sum_{p \in P} \int_{\Lambda^{N}} \left( \prod_{k=1}^{N} 1(x_k \in \Lambda_{up(k)}) dx_k \right) |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2
\]

(2.34)

\[
= \sum_{p \in P} \sum_{\lambda} \int_{\Lambda^{N-|p^{-1}(\lambda)|}} \left( \prod_{k:p(k) \neq \lambda} 1(x_k \in \Lambda_{up(k)}) dx_k \right) \times \left( \int_{\Lambda^{p^{-1}(\lambda)}} \sum_{j:p(j) = \lambda} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 \prod_{k:p(k) = \lambda} dx_k \right)
\]

(2.35)

Here, the last integral of the right hand side of (2.34) represents the kinetic energy in \( \Lambda_{u_\lambda} \). Likewise, we also have

\[
\sum_{i \neq j}^{N} \int_{\Lambda^{N-2}} dx_1 dx_2 \cdots \widehat{dx_i} \cdots dx_j \cdots dx_N \left( \sum_{\lambda} \int_{\Lambda_{u_\lambda}} \bar{q}_{ij} |\Psi(x)|^2 dx_i dx_j \right)
\]

(2.36)

Then, applying Lemma 2.6 from (2.32), (2.34), and (2.35) we get

\[
\epsilon \langle \Psi, H_N \Psi \rangle + \frac{1}{2} \sum_{j=1}^{N} \int_{\Lambda^{N}} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 dx
\]

\[
+ (1 - \epsilon) \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^{3} \sum_{i \neq j}^{N} \int_{\Lambda^{N-2}} dx_1 dx_2 \cdots \widehat{dx_i} \cdots dx_j \cdots dx_N \left( \sum_{\lambda} \int_{\Lambda_{u_\lambda}} \bar{q}_{ij} |\Psi(x)|^2 dx_i dx_j \right)
\]

(2.37)

\[
\geq (1 - Ce) \frac{4\pi a}{\ell_1} \int_{\Lambda^{N}} \sum_{\lambda} f(\mathcal{N}(\Lambda_{u_\lambda})) |\Psi(x)|^2 dx,
\]

where \( f \) is defined as in (2.29). Note that in (2.37) we only used half of kinetic energy. Combining (2.31), (2.32), and (2.37), we get

\[
\langle \Psi, H_N \Psi \rangle \geq \left( \frac{1}{2} - \epsilon \right) \sum_{j=1}^{N} \int_{\Lambda^{N}} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 dx + (1 - Ce) \frac{4\pi a}{\ell_1} \int_{\Lambda^{N}} \sum_{\lambda} f(\mathcal{N}(\Lambda_{u_\lambda})) |\Psi(x)|^2 dx.
\]

(2.38)

Now we consider larger cubic boxes of side length \( \ell_2 \). We let \( \Lambda_{u_\theta} \) be a cubic box of side length \( \ell_2 \), which is a union of \( 2^{3\theta} \) \( \Lambda_{u_\lambda} \)'s, where \( \theta \) is an index for those larger cubic boxes. With a shorthand notation
\( x = (x_1, x_2, \ldots, x_N) \), we obtain from Lemma 2.7 that
\[
\frac{1}{2} - \varepsilon \sum_{j=1}^{N} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 dx + (1 - C \varepsilon) \frac{4 \pi a}{\ell_3^2} \int_{\Lambda} \sum_{\lambda} f(\mathcal{N}(\Lambda_{\lambda})) |\Psi(x)|^2 dx
\]
\[
\geq (1 - C \varepsilon) \frac{4 \pi a}{\ell_3^2} \int_{\Lambda} \sum_{\theta} \tilde{f}(\mathcal{N}(\Lambda_{\theta})) |\Psi(x)|^2 dx,
\]
(2.39)

where \( \tilde{f} \) is a convex function such that \( \tilde{f}(n) = n(n - 1) \) if \( n \leq \rho \ell_3^2 \). So far we have proved that
\[
\langle \Psi, H_N \Psi \rangle \geq \inf_{u} (1 - C \varepsilon) \frac{4 \pi a}{\ell_3^2} \int_{\Lambda} \sum_{\theta} \tilde{f}(\mathcal{N}(\Lambda_{\theta})) |\Psi(x)|^2 dx.
\]
(2.40)

We are left to minimize \( \sum_{\theta} \tilde{f}(\mathcal{N}(\Lambda_{\theta})) \). Since \( \tilde{f} \) is convex and \( \sum_{\theta} \mathcal{N}(\Lambda_{\theta}) = N \), we can use Jensen’s inequality to get, for any \( x_1, x_2, \ldots, x_N \),
\[
\sum_{\theta} \tilde{f}(\mathcal{N}(\Lambda_{\theta})) \geq \frac{L_3^3}{\ell_3^2} \tilde{f}(\rho \ell_3^2) \geq \frac{L_3^3}{\ell_3^2} \rho \ell_3^3 (\rho \ell_3^2 - 1) \geq (1 - C \rho \ell_3^2 \log \rho^2) N \rho \ell_3^3.
\]
(2.41)

Note that this lower bound is independent of \( u \). Hence
\[
\frac{\langle \Psi, H_N \Psi \rangle}{\langle \Psi, \Psi \rangle} \geq (1 - C \varepsilon) \frac{4 \pi a}{\ell_3^2} N \rho \ell_3^3 = 4 \pi a \rho N (1 - C \varepsilon).
\]
(2.42)

This shows that, when \( \rho \) is small enough, there exists a constant \( C_0 \) such that
\[
\lim_{N, L \to \infty} \frac{E(N, \Lambda)}{N} \geq 4 \pi a \rho (1 - C_0 \varepsilon),
\]
(2.43)

which was to be proved.

\( \square \)

### 3. Lower bound estimates

In this section, we prove Lemma 2.5, Lemma 2.6, and Lemma 2.8, which were used in the proof of the main theorem.

#### 3.1. Conversion into a soft potential

**Proof of Lemma 2.5** Fix \( j \) and consider \( x_1, x_2, \ldots, \hat{x}_j, \ldots, x_N \) to be fixed. From that \( \Phi_j = W_j^{-1} \Psi \), we get
\[
\int_{\Lambda} |\nabla_j \Psi(x)|^2 dx + \frac{1}{2} \sum_{i \neq j} \int_{\Lambda} V(x_i - x_j) |\Psi(x)|^2 dx_j
\]
\[
= \int_{\Lambda} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 dx_j + \sum_{i \neq j} \int_{\Lambda} \left[ W_j(x) \left( - \Delta_j + \frac{1}{2} V(x_i - x_j) W_j(x) \right) \right] |\Phi_j(x)|^2 dx_j
\]
(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 \( \ell_0 \) centered at \( x_i \). When \( x_j \in B(x_i, \ell_0) \), \( W_j(x) = 1 - \tau(\ell_0, x_i - x_j) \), thus
\[
\left( - \Delta_j + \frac{1}{2} V(x_i - x_j) \right) W_j(x) = \ell_0(\ell_0) W_j(x).
\]
(3.2)
When \( x_j \notin B(x_i, \ell_0) \), \((-\Delta_j + \frac{1}{2} V(x_i - x_j)) W_j(x) = 0 \). Thus,
\[
\int_{\Lambda} \left[ W_j(x) \left( -\Delta_j + \frac{1}{2} V(x_i - x_j) \right) W_j(x) \right] |\Phi_j(x)|^2 dx_j
= \int_{\Lambda} c_0(\ell_0) 1(|x_i - x_j| \leq \ell_0) |W_j(x)|^2 |\Phi_j(x)|^2 dx_j = \int_{\Lambda} F_{ij}(x) q(\ell_0, x_i - x_j) |\Psi(x)|^2 dx_j. \tag{3.3}
\]

(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 \( x_j \in B(x_i, t_{ij}) \), \( W_i(x) = 1 - \tau(t_{ij}, x_i - x_j) \), thus
\[
(-\Delta_j + \frac{1}{2} V(x_i - x_j)) W_j(x) = c_0(t_{ij}) W_j(x). \tag{3.4}
\]
When \( x_j \notin B(x_i, t_{ij}) \), \( W_j(x)(-\Delta_j + \frac{1}{2} V(x_i - x_j)) W_j(x) = 0 \). Thus,
\[
\int_{\Lambda} \left[ W_j(x) \left( -\Delta_j + \frac{1}{2} V(x_i - x_j) \right) W_j(x) \right] |\Phi_j(x)|^2 dx_j = \int_{\Lambda} e_0(t_{ij}) 1(|x_i - x_j| \leq t_{ij}) |W_j(x)|^2 |\Phi_j(x)|^2 dx_j = \int_{\Lambda} (1 - F_{ij}(x)) G_{ij}(x) q(t_{ij}, x_i - x_j) |\Psi(x)|^2 dx_j. \tag{3.5}
\]

(3) When \( F_{ij} = G_{ij} = 0 \), we have that \( W_j(x) = 1 \), thus we get
\[
\int_{\Lambda} \left[ W_j(x) \left( -\Delta_j + \frac{1}{2} V(x_i - x_j) \right) W_j(x) \right] |\Phi_j(x)|^2 dx_j \geq 0. \tag{3.6}
\]

From cases (1)-(3), \[3.1\] implies
\[
\int_{\Lambda} \left( |\nabla_j \Psi(x_j)|^2 + \frac{1}{2} \sum_{i:j \neq j} V(x_i - x_j) |\Psi(x_j)|^2 \right) dx_j \\
\geq \int_{\Lambda} |W_j(x)|^2 |\nabla_j \Phi_j(x)|^2 dx_j + \sum_{i:j \neq j} \int_{\Lambda} \left[ F_{ij}(x) q(\ell_0, x_i - x_j) + (1 - F_{ij}(x)) G_{ij}(x) q(t_{ij}, x_i - x_j) \right] |\Psi(x)|^2 dx_j. \tag{3.7}
\]

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

3.2. Decomposition of \( \Lambda \).

**Proof of Lemma \[2.2\]** Recall that we let \( u \in [-\ell_1/2, \ell_1/2]^3 = \Gamma, G = \ell_1 \mathbb{Z}^3 \cap \Lambda \), and for \( \lambda \in G \), \( \Lambda_{u\lambda} \) be a cubic box of side length \( \ell_1 \) centered at \((u + \lambda)\). Here, \( u \) corresponds to the origin of the grid that divides \( \Lambda \) into small boxes \( \Lambda_{u\lambda} \). Note that the positions of those boxes depend on \( u \).

Define \( \tilde{\chi}_{u\lambda} \) by
\[
\tilde{\chi}(x) := \begin{cases} 
1 & \text{if } x \in [-\frac{\ell_1}{2} + 2\ell_0, \frac{\ell_1}{2} - 2\ell_0]^3 \\
0 & \text{otherwise}
\end{cases} \tag{3.8}
\]
and \( \tilde{\chi}_{u\lambda} := \tilde{\chi}(x - u - \lambda) \). If \( \tilde{\chi}_{u\lambda}(x_i) = 1 \), 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
\[
\frac{1}{|\Gamma|} \int_{\Gamma} du \sum_{\lambda \in G} \tilde{\chi}_{u\lambda}(x) = \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \tag{3.9}
\]
for any \( x \in \Lambda \). This means that the probability of having \( x \in \Lambda \) to satisfy \( \tilde{\chi}_{u\lambda} = 1 \) for a \( \lambda \) is \( (\ell_1 - 4\ell_0)^3/\ell_1^3 \).

From the definitions of \( \tilde{q}_{ij} \) and \( \tilde{\chi}_{u\lambda} \), we have
\[
\tilde{q}_{ij}(x) = \sum_{\lambda \in G} \tilde{\chi}_{u\lambda}(x_i) \tilde{q}_{ij}(x). \tag{3.10}
\]
(\tilde{q}_{ij} \text{ depends on } u, \text{ but we omit it.}) Thus, for any \( \Psi(x) \in L^2(\Lambda^N) \),
\[
\frac{1}{|I|} \int_{\Lambda} du \int_{\Lambda^2} \tilde{q}_{ij} |\Psi(x)|^2 dx_{i}dx_{j} = \frac{1}{|I|} \int_{\Lambda} du \sum_{\lambda \in \Theta} \chi_{u\lambda}(x_{i})q_{ij}|\Psi(x)|^2 dx_{i}dx_{j}
\]
\[
= \left( \frac{\ell_1 - 4\ell_0}{\ell_1} \right)^3 \int_{\Lambda^2} q_{ij}|\Psi(x)|^2 dx_{i}dx_{j}. \tag{3.11}
\]
Integrating (3.11) with respect to \( dx_{1}dx_{2} \cdots dx_{i} \cdots dx_{j} \cdots dx_{N} \) and summing over \( i, j \) gives
\[
\sum_{i \neq j}^{N} \int_{\Lambda^N} q_{ij}|\Psi(x)|^2 dx = \frac{1}{|I|} \int_{\Lambda} du (\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i \neq j}^{N} \tilde{q}_{ij}|\Psi(x)|^2 dx \geq \inf_{u}(\frac{\ell_1}{\ell_1 - 4\ell_0})^3 \sum_{i \neq j}^{N} \tilde{q}_{ij}|\Psi(x)|^2 dx. \tag{3.12}
\]
This proves the lemma. \qed

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 \( n \geq 2 \). Define an operator

   \[
   T_j = -W_j^{-1}\nabla_j W_j^2 \nabla_j W_j^{-1}
   \tag{3.13}
   \]

   on the functions in \( L^2(B^n) \) with \( W_j \)-Neumann boundary conditions in \( B \) in the sense that, for a function \( \psi \in L^2(B^n) \),

   \[
   T_j \psi = -W_j^{-1}\nabla_j (W_j^2 \nabla_j (W_j^{-1} \psi)).
   \tag{3.14}
   \]

   For \( i, j \in \{1, 2, \cdots, n\}, i \neq j \), we want to consider an operator

   \[
   \frac{T_i}{4(n-1)} + \frac{T_j}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \tilde{q}_{ij},
   \tag{3.15}
   \]

   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

   \[
   \frac{T_j}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \tilde{q}_{ij}
   \tag{3.16}
   \]

   by Temple’s inequality [19], 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

   \[
   \int_{B} W_j \psi(x_j) dx_j = 0,
   \tag{3.17}
   \]

   we apply Poincare’s inequality to obtain

   \[
   \int_{B} |\psi(x_j) - \bar{\psi}|^2 dx_j \leq C|B|^2 \int_{B} |\nabla \psi(x_j)|^2 dx_j,
   \tag{3.18}
   \]

   where

   \[
   \bar{\psi} = \frac{1}{|B|} \int_{B} \psi(x_j) dx_j.
   \tag{3.19}
   \]
We also have
\[
\int_B |\psi(x_j) - \tilde{\psi}|^2 dx_j = \int_B |\psi(x_j)|^2 dx_j - \frac{1}{|B|} \left( \int_B |\psi(x_j)| dx_j \right)^2
\]
\[
= \int_B |\psi(x_j)|^2 dx_j - \frac{1}{|B|} \left( \int_B (1 - W_j) |\psi(x_j)| dx_j \right)^2 \geq \int_B |\psi(x_j)|^2 dx_j - \frac{c_0^2}{|B|} \left( \int_B |\psi(x_j)| dx_j \right)^2
\]
\[
\geq (1 - c_0^2) \int_B |\psi(x_j)|^2 dx_j.
\]
Thus, we can see that, in a box of side length \(\ell_1\),
\[
\text{(gap of } T_j) \geq C\ell_1^{-2}. \tag{3.20}
\]

In order to use Temple’s inequality, we first need to check that
\[
\frac{T_j}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \tilde{q}_{ij} \geq 0,
\]
which is obvious since \(T_j \geq 0\) and \(\tilde{q}_{ij} \geq 0\). Let \(\langle F \rangle_{W_i}\) denotes
\[
\langle F \rangle_{W_i} = \int_B dx_i F|W_i|^2 / \int_B dx_i |W_i|^2.
\]
We also need to have
\[
\text{(gap of } \frac{T_j}{4(n-1)}) \gg \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \langle \tilde{q}_{ij} \rangle_{W_j},
\]
which can be easily seen from that
\[
\text{(gap of } \frac{T_j}{4(n-1)}) \geq Cn^{-1}\ell_1^{-2} \geq C\rho^\frac{2}{3} |\log \rho|^{\frac{2}{3} - 5\eta}
\]
\[
\gg C\rho^\frac{2}{3} |\log \rho|^{-2} \geq C\|q_{ij}\|_{\infty} \geq \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \langle \tilde{q}_{ij} \rangle_{W_j}. \tag{3.25}
\]
Thus, we can indeed use Temple’s inequality to obtain that
\[
\frac{T_j}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \tilde{q}_{ij} \geq \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \langle \tilde{q}_{ij} \rangle_{W_j} - C \frac{\langle \tilde{q}_{ij} \rangle_{W_j}^2 - \langle \tilde{q}_{ij} \rangle_{W_j}^2}{n^{-1}\ell_1^2}.
\]

We want to estimate \(\langle \tilde{q}_{ij} \rangle_{W_j}\) and \(\langle \tilde{q}_{ij}^2 \rangle_{W_j}\). It follows from Lemma 5.1 that, for all \(\ell_{-1} \leq \kappa \leq \ell_0\),
\[
\tau(\kappa, x_i - x_j) \leq \frac{C}{|x_i - x_j|} \tag{3.26}
\]
Hence,
\[
\tilde{q}_{ij}|W_j|^2 \geq \tilde{q}_{ij}(1 - \frac{C}{|x_i - x_j|})^2 \geq \tilde{q}_{ij}(1 - \frac{C}{|x_i - x_j|}). \tag{3.27}
\]

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 \(x_1, x_2, \ldots, x_i, \ldots, x_j, \ldots, x_N\), i.e.,
\[
S_{-1} = \{ x \in B : \forall k \neq i, j, |x - x_k| > 2\ell_{-1} \},
\]
\[
S_0 = \{ x \in B : \forall k \neq i, j, |x - x_k| > 2\ell_0 \}.
\]
Let \(\tilde{B}\) be a set of all points in \(B\) that are not within distance \(2\ell_0\) to the boundary of \(B\), i.e.,
\[
\tilde{B} = \{ x \in B : d(x, \partial B) \geq 2\ell_0 \}.
\]
By definition, \(\tilde{q}_{ij} = 0\) if \(x_i \in B \setminus S_{-1}\) or \(x_i \in B \setminus \tilde{B}\), and \(\tilde{q}_{ij} = q_{ij}\) if \(x_i \in S_{-1} \cap \tilde{B}\).
When \( x_i \in S_0 \cap \bar{B} \),
\[
\int_B \tilde{q}_{ij} |W_j|^2 \, dx_j \geq \int_B \tilde{q}_{ij} (1 - \frac{C}{|x_i - x_j|}) \, dx_j \geq e_0(\ell_0) \int_{|x_i - x_j| \leq \ell_0} (1 - \frac{C}{|x_i - x_j|}) \, dx_j
\]
\[
\geq (1 - \frac{C}{\ell_0}) e_0(\ell_0) \int_{|x_i - x_j| \leq \ell_0} 1 \, dx_j \geq 4\pi a (1 - \frac{C}{\ell_0}).
\]  
(3.32)

When \( x_i \in (S_1 \setminus S_0) \cap \bar{B} \),
\[
\int_B \tilde{q}_{ij} |W_j|^2 \, dx_j \geq \int_B \tilde{q}_{ij} (1 - \frac{C}{|x_i - x_j|}) \, dx_j \geq e_0(t_{ij}) \int_{|x_i - x_j| \leq t_{ij}} (1 - \frac{C}{|x_i - x_j|}) \, dx_j
\]
\[
\geq (1 - \frac{C}{t_{ij}}) e_0(t_{ij}) \int_{|x_i - x_j| \leq t_{ij}} 1 \, dx_j \geq 4\pi a (1 - \frac{C}{t_{ij}}).
\]  
(3.33)

For \( \langle \tilde{q}_{ij}^{-2} \rangle_{W_j} \), when \( x_i \in S_0 \cap \bar{B} \),
\[
\int_B \tilde{q}_{ij}^{-2} |W_j|^2 \, dx_j \leq \int_B \tilde{q}_{ij}^{-2} \, dx_j \leq C\ell_0^{-3},
\]  
(3.34)

and, when \( x_i \in (S_1 \setminus S_0) \cap \bar{B} \),
\[
\int_B \tilde{q}_{ij}^{-2} |W_j|^2 \, dx_j \leq \int_B \tilde{q}_{ij}^{-2} \, dx_j \leq C t_{ij}^{-3}.
\]  
(3.35)

Since we know from Lemma 5.2 that
\[
(1 - Cn\ell_0^2 \ell_1^{-1}) \leq \int_B |W_j|^2 \, dx_j \leq \ell_1^3,
\]  
(3.36)

from (3.32) and (3.33), we get
\[
\left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \langle \tilde{q}_{ij} \rangle_{W_j} = \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \int_B \tilde{q}_{ij} |W_j|^2 \, dx_j / \int_B |W_j|^2 \, dx_j
\]
\[
\geq \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \frac{4\pi a}{\ell_1^3} \left( 1 - \frac{C}{\ell_0} \right) \cdot 1(x_i \in S_0 \cap \bar{B}) + (1 - \frac{C}{t_{ij}}) \cdot 1(x_i \in (S_1 \setminus S_0) \cap \bar{B}) \right),
\]  
(3.37)

and, from (3.34) and (3.35), we get
\[
\langle \tilde{q}_{ij}^{-2} \rangle_{W_j} - \langle \tilde{q}_{ij}^{-2} \rangle_{W_{ij}} \leq \frac{\langle \tilde{q}_{ij}^{-2} \rangle_{W_j}}{n^{-1}\ell_1^{-3}} = n\ell_1^2 \int_B \tilde{q}_{ij}^{-2} |W_j|^2 \, dx_j / \int_B |W_j|^2 \, dx_j
\]
\[
\leq C \frac{n\ell_1^2}{\ell_0^2} \ell_1^{-3} \cdot 1(x_i \in S_0 \cap \bar{B}) + C \frac{n\ell_1^2}{t_{ij}^3} \ell_1^{-3} \cdot 1(x_i \in (S_1 \setminus S_0) \cap \bar{B}).
\]  
(3.38)

Inserting (3.37) and (3.38) into (3.36), we obtain that
\[
\frac{T_j}{4(n - 1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \langle \tilde{q}_{ij} \rangle
\]
\[
\geq \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \frac{4\pi a}{\ell_1^3} \left( 1 - \frac{C}{\ell_0} - C \frac{n\ell_1^2}{\ell_0^2} \right) \cdot 1(x_i \in S_0 \cap \bar{B}) + (1 - \frac{C}{t_{ij}} - C \frac{n\ell_1^2}{t_{ij}^3}) \cdot 1(x_i \in (S_1 \setminus S_0) \cap \bar{B})
\]
\[
\geq \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \frac{4\pi a}{\ell_1^3} \left( 1 - \frac{C}{\ell_0} - C \frac{n\ell_1^2}{\ell_0^2} \right) \cdot 1(x_i \in S_0 \cap \bar{B}) + (1 - \frac{C}{t_{ij}} - C \frac{n\ell_1^2}{t_{ij}^3}) \cdot 1(x_i \in (S_1 \setminus S_0) \cap \bar{B})
\]  
(3.39)

Here, for the last inequality, we used \( n\ell_1^2 \geq \ell_1^3 \gg \ell_0^3 \) and, when \( x_i \in (S_1 \setminus S_0) \cap \bar{B} \), \( n\ell_1^2 \geq \ell_1^3 \gg \ell_0^3 \gg \ell_{ij}^2 \).
Let
\[ \xi_i(x_n) = \frac{4\pi a}{\ell_1} \left( 1 - C \frac{\rho\ell_1^2}{\ell_0^2} \right) 1(x_i \in S_0 \cap \tilde{B}) + \left( 1 - C \frac{\rho\ell_1^2}{\ell_{ij}^2} \right) 1(x_i \in (S_{-1} \setminus S_0) \cap \tilde{B}) \right). \] (3.40)

Note that \( S_0 \) and \( S_{-1} \) are independent of \( x_i \) and \( x_j \), and \( \xi_i \) is independent of \( x_j \). We apply Temple's inequality to
\[ \frac{T_i}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \xi_i \] (3.41)
with \( [4(n-1)]^{-1}T_i \) as the unperturbed part. Then, we get
\[ \frac{T_i}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \xi_i \geq \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 (\xi_i)_W - C \frac{(\xi_i^2)_W}{n^{-1} \ell_1} \] (3.42)

We now estimate \( (\xi_i)_W \). By definition,
\[ (\xi_i)_W \geq \ell_1^{-3} \int_{S_{-1} \cap \tilde{B}} \xi_i W_i^2 dx_i \]
\[ \geq \frac{4\pi a}{\ell_1^3} (1 - C \frac{\rho\ell_1^2}{\ell_0^2}) \int_{S_{-1} \cap \tilde{B}} \ell_1^{-3} |W_i|^2 dx_i - C \int_{S_{-1} \cap \tilde{B}} \frac{\ell_1^{-1} \rho\ell_1^2}{\ell_{ij}^2} dx_i. \] (3.43)
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 \( x_i \) and (2) the union of annuli \( \{x : 2\ell_{-1} \leq |x - x_k| \leq 2\ell_0\} \) for all \( x_k \in B \) other than \( x_i \) and \( x_j \) covers \( (S_{-1} \setminus S_0) \). Thus,
\[ \int_{(S_{-1} \cap \tilde{B}) \setminus \tilde{B}} t_{ij}^3 dx_i \leq \sum_{k \neq i} \int_{2\ell_{-1} \leq |x_i - x_k| \leq 2\ell_0} |x_i - x_k|^{-3} dx_i. \] (3.44)
To estimate the first term in the right hand side of (3.43), we use
\[ \int_{S_{-1} \cap \tilde{B}} |W_i|^2 dx_i \geq \int_{S_{-1} \cap \tilde{B}} dx_i - \int_B (1 - |W_i|^2) dx_i \geq |S_{-1} \cap \tilde{B}| - Cn\ell_0^2 \] (3.45)
where the last inequality follows from Lemma 5.2. Since \( |S_{-1}| \geq \ell_1^3 - Cn(\ell_{-1})^3 \) and \( |B| = (\ell_1 - 4\ell_0)^3 \), from (3.43), (3.44), and (3.45), we get
\[ (\xi_i)_W \]
\[ \geq \frac{|S_{-1} \cap \tilde{B}|}{\ell_1^3} \frac{4\pi a}{\ell_1^4} (1 - C \frac{\rho\ell_1^2}{\ell_0^2}) (1 - Cn\ell_0^2) - Cn \sum_{k \neq i} \int_{2\ell_{-1} \leq |x_i - x_k| \leq 2\ell_0} \ell_{ij}^{-4} |x_i - x_k|^{-3} dx_i \]
\[ \geq \left( \frac{\ell_1 - 4\ell_0}{\ell_1} \right)^3 \frac{4\pi a}{\ell_1^6} (1 - Cn(\ell_{-1})^3) - Cn\ell_0^2 \frac{\ell_0^2}{\ell_1^6} - Cn^2 \frac{|\log \rho|}{\ell_1^6}. \] (3.46)
Since \( (\xi_i^2)_W \), \( \|\xi_i\|_\infty \leq C\ell_1^{-6} \),
\[ \frac{1}{n^{-1} \ell_1^2} \leq Cn\ell_1^{-4}. \] (3.47)
Thus, combining (3.39), (3.42), (3.46), and (3.47), we obtain that
\[ \frac{T_i}{4(n-1)} + \frac{T_j}{4(n-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 q_{ij} \]
\[ \geq (1 - Cn(\ell_{-1})^3 - Cn \frac{\rho\ell_0^2}{\ell_1^4} - Cn \frac{\rho\ell_1^2}{\ell_1^4} - Cn^2 |\log \rho|) \frac{4\pi a}{\ell_1^4}. \] (3.48)
This implies, when \( x_{n+1}, \ldots, x_N \) are outside of \( B \), for any \( \psi(x_n) \in L^2(B^n) \) with \( \phi_l = W_l^{-1} \psi \) for \( l = 1, 2, \ldots, n, \)

\[
\frac{1}{4(n-1)} \int_{B^n} (|W_l|^2 \nabla \phi_l(x_n)|^2 + |W_l|^2 |\nabla \phi_l(x_n)|^2) dx_n + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \int_{B^n} q_{ij} |\psi(x_n)|^2 dx_n \\
\geq (1 - C\epsilon) \frac{4\pi a}{\ell_1^4} \|\psi\|_2^2. \tag{3.49}
\]

This shows how we can get the lower bound for the ground state energy in a soft potential regime. \( q_{ij} \) 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

\[
\frac{1}{2} \sum_{j=1}^n \int_{B^n} |W_l|^2 |\nabla \phi_l(x_n)|^2 dx_n + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \sum_{i \neq j} \int_{B^n} \tilde{q}_{ij} |\psi(x_n)|^2 dx_n \\
\geq n(n-1)(1 - C\epsilon) \frac{4\pi a}{\ell_1^4} \|\psi\|_2^2. \tag{3.50}
\]

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

Let \( p := 9\rho^3 \). Here, \( n \) satisfies

\[
n^{-\frac{(\ell_1 - 3)^3}{n}} \ll 1, \quad np \left| \frac{\log p}{\ell_1} \right| \ll 1, \quad n^2 \ell_0^2 \ell_1 \ll 1. \tag{3.51}
\]

We form particle groups in \( B \), each of which contains at most \( p \) particles, such that \( G_1 = \{x_1, x_2, \ldots, x_p\} \), \( G_2 = \{x_{p+1}, \ldots, x_{2p}\} \), \ldots, \( G_{\lfloor n/p \rfloor} = \{x_{(n/p)-1}+1, \ldots, x_{n/p}\} \), \( G_{\lfloor n/p \rfloor+1} = \{x_{n/p}+1, \ldots, x_n\} \). Note that the number of groups with \( p \) particles is \( \lfloor n/p \rfloor \) and \( n/p \geq n/2p \).

For \( i, j \in G_k \), we consider

\[
\frac{T_i}{4(p-1)} + \frac{T_j}{4(p-1)} + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \tilde{q}_{ij} \tag{3.52}
\]

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

\[
\frac{1}{4(p-1)} \int_{B^n} (|W_l|^2 |\nabla \phi_l(x_n)|^2 + |W_l|^2 |\nabla \phi_l(x_n)|^2) dx_n + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \int_{B^n} \tilde{q}_{ij} |\psi(x_n)|^2 dx_n \\
\geq (1 - Cn^{-\frac{(\ell_1 - 3)^3}{n}} - C_{p^n} \left| \frac{\log p}{\ell_1} \right| \frac{4\pi a}{\ell_1^4} \|\psi\|_2^2. \tag{3.53}
\]

We apply this inequality to all particles in the particle group \( G_1 \), then to all particle groups in \( B \). Using (3.51), we obtain

\[
\frac{1}{2} \sum_{j=1}^n \int_{B^n} |W_l|^2 |\nabla \phi_l(x_n)|^2 dx_n + \left( \frac{\ell_1}{\ell_1 - 4\ell_0} \right)^3 \sum_{i \neq j} \int_{B^n} \tilde{q}_{ij} |\psi(x_n)|^2 dx_n \\
\geq \left( \frac{8}{9} \right)^\frac{4\pi a}{\ell_1^4} \|\psi\|_2^2 \geq (4\rho^3 - 1)n(1 - C\epsilon) \frac{4\pi a}{\ell_1^4} \|\psi\|_2^2. \tag{3.54}
\]

We note that, while in Lemma 2.6 the two cases \( n \leq 2\rho^3 \) and \( n > 2\rho^3 \) are distinguished, we choose the size of \( p \) to be \( 9\rho^3 \). If we would choose the size of \( p \) to be \( 2\rho^3 \), we would get a lower bound that is not convex at \( n = 2\rho^3 \), since the lower bound would increase proportionally to
n. We need a space to connect two different formulas so that the resulting lower bound becomes convex, thus the distinction is between \( n \leq 9\rho \ell_1^3 \) and \( n > 9\rho \ell_1^3 \).

(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 \( \ell_1 \),

\[
E(n, B) \geq 4\pi a(\frac{n}{\ell_1^3})n(1 - C(\frac{n}{\ell_1^3})^{\frac{1}{17}}) \tag{3.55}
\]

when the density \( (n/\ell_1^3) \) is sufficiently small and

\[
\ell_1 \geq C(\frac{n}{\ell_1^3})^{-\frac{1}{17}}. \tag{3.56}
\]

(See Theorem 2.4 in [15].)

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

\[
\ell_1 \geq C\rho^{-\frac{1}{3}}|\log \rho|^{-\frac{1}{3}} \tag{3.57}
\]

and the density in this case

\[
m = 9\epsilon^{-1}\rho = 9\rho^{\frac{1}{3}}|\log \rho|^{-\frac{2}{3}} \rightarrow 0 \tag{3.58}
\]

as \( \rho \to 0 \). Thus we can indeed use (3.55) to obtain that

\[
E(m, B) \geq 4\pi a\frac{m^2}{\ell_1^3}(1 - C(\frac{2m}{\ell_1^3})^{\frac{1}{17}}) \tag{3.59}
\]

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,

\[
E(n, B) \geq \frac{n}{m} E(m, B) \geq (4\rho \ell_1^3)\frac{4\pi a}{\ell_1^3}(1 - C\epsilon)n(1 - C\epsilon). \tag{3.60}
\]

Since, \( n \) falls into one of the above categories, we get the desired lemma.

\( \square \)

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 \( \Lambda_{\ell_2} \), a box of side length \( \ell_2 \) is given. We note:

\[
\ell_1 \ll \ell_2 \sim \rho^{-\frac{1}{3}}|\log \rho|^{-\frac{1}{3}}. \tag{4.1}
\]

More specifically, for a box \( B \) with a side length between \( \ell_1 \) and \( \ell_2 \), 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 \( \ell_2 \), 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 \( \Lambda_A \) and \( \Lambda_B \) 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 \( \Lambda_A \) and \( \Lambda_B \). We assume the
potential energy in $\Lambda_A$ and $\Lambda_B$ 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

$$H_\alpha = \alpha \sum_{i=1}^n (-\Delta_i) + n_A(n_A - 1)/|\Lambda_A| + n_B(n_B - 1)/|\Lambda_B|$$  \hspace{1cm} (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 $-\alpha \Delta_1$, a small potion of $T_1$ 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 $\ell_1$. We consider $(2^{3h}/2)$ pairs of adjacent boxes, and for each pair that consists of two adjacent boxes of same size $\Lambda_A$ and $\Lambda_B$. 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 $T_1$’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 $\rho$).

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 $\rho$).

We keep repeating it until the side length of a box becomes $\ell_2$, 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 $1 \leq s \leq 3h$, where $\ell_2 = 2^h \ell_1$. We let

$$\ell(s) := 2^{\frac{1}{2^n+1}} \ell_1.$$  \hspace{1cm} (4.4)

$s$ satisfies $\rho^{-\frac{4}{7}} |\log \rho|^{-\frac{4}{7}+n} \sim \ell_1 \leq \ell(s) \leq \ell_2 \sim \rho^{-\frac{4}{7}} |\log \rho|^{-\frac{4}{7}}$, 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 $s = 3h$.

- $\Lambda_A$ and $\Lambda_B$ are boxes such that the volume of each box $|\Lambda_A| = |\Lambda_B| = 2^{s-1} \ell_1^3$ and the dimensions of $\Lambda_A$, $\Lambda_B$, 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 $\Lambda_A$ and $\Lambda_B$, respectively, when $x_{n+1}, x_{n+2}, \cdots, x_N$ are outside $(\Lambda_A \cup \Lambda_B)$, i.e.,

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

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

$\Lambda_A$ are $\Lambda_B$ 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 \((1 \leq s \leq 3h + 1)\)
  \[
  f_s(t) := \begin{cases} 
  t(t-1) & \text{if } t \leq K_s \\
  (2K_s - 1)t - K_s^2 & \text{if } t > K_s
  \end{cases} 
  \] (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 \(K_1 := 2\rho \ell_1^3\) (4.7) and we choose \(K_s\) such that
  \[
  2K_s - K_{s+1} \gg |\log \rho|^{1/2} \sqrt{K_{s+1}} \quad \text{and} \quad K_s > 2^{s-1} \rho \ell_1^3. 
  \] (4.8)

For example,

\[
K_s = (2 - (1 - \frac{1}{2^{s-1}})|\log \rho|^{-\eta}) \cdot 2^{s-1} \rho \ell_1^3. 
\] (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 (2.20), 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 \(\rho\). (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 \(x_{n+1}, x_{n+2}, \ldots, x_N\) are fixed outside \((\Lambda_A \cup \Lambda_B)\). Then, for any \(\psi(x_n) \in L^2((\Lambda_A \cup \Lambda_B)^n)\) with \(\phi_j = W_j^{-1} \psi\) for \(j = 1, 2, \ldots, n\) (Here \(x_n = (x_1, x_2, \ldots, x_n)\)),

\[
\frac{(4\pi a)^{-1}}{|\log \rho|} \sum_{j=1}^n \left( \int_{(\Lambda_A \cup \Lambda_B)^n} |W_j|^2 |\nabla_j \phi_j(x_n)|^2 dx_n + \int_{(\Lambda_A \cup \Lambda_B)^n} \frac{f_s(M(A)) + f_s(M(B))}{|\Lambda|} |\psi(x_n)|^2 dx_n \right) 
\geq (1 - C \rho^{\frac{1}{4}} \int_{(\Lambda_A \cup \Lambda_B)^n} \frac{f_{s+1}(n)}{|\Lambda_A \cup \Lambda_B|} |\psi(x_n)|^2 dx_n. 
\] (4.10)

Here, the constant \(C\) does not depend on \(s\).

Note that the factor \((4\pi a)^{-1}\) can be a general constant \(C\) in this proposition and subsequent lemmas. We keep this factor, however, in order to use it in the proof of Lemma 2.7.

Proposition 4.1 shows the outcome of box doubling method when it is applied to \(\Lambda_A\) and \(\Lambda_B\) 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 \(\Lambda_{\ell_2}\) whose side length is \(\ell_2\) and \(B_1, B_2, \ldots, B_{2^h}\) are small cubic cells of side length \(\ell_1\) such that \(\bigcup_{k=1}^{2^h} B_k = \Lambda_{\ell_2}\). We have that \(x_{n+1}, x_{n+2}, \ldots, x_N\) are fixed outside \(\Lambda_{\ell_2}\) and we let \(x_n = (x_1, x_2, \ldots, x_n)\). We want to prove that, for a given \(\psi(x_n) \in L^2(\Lambda_{\ell_2}^n)\) with \(\phi_j = W_j^{-1} \psi\) for \(j = 1, 2, \ldots, n\),

\[
\left( \frac{1}{2} - \epsilon \right) \sum_{j=1}^n \int_{\Lambda_{\ell_2}} |W_j|^2 |\nabla_j \phi_j(x_n)|^2 dx_n + \frac{4\pi a}{\ell_1} \sum_{k=1}^{2^h} \int_{\Lambda_{\ell_2}} f(N(B_k)) |\psi(x_n)|^2 dx_n 
\geq \tilde{f}(n)(1 - C \rho^{\frac{1}{4}} |\log \rho|) \frac{4\pi a}{\ell_2^3} \|\psi\|_2^2. 
\] (4.11)
Here,

\[
\bar{f}(t) := \begin{cases} 
 t(t-1) & \text{if } t \leq \rho \ell^3_2 \\
 (2\rho \ell^3_2 - t) - (\rho \ell^3_2)^2 & \text{if } t > \rho \ell^3_2.
\end{cases}
\]  

(4.12)

Since

\[
3h = \frac{\log(\ell^3_2/\ell^3_1)}{\log 2} \leq \left(\frac{1}{2} - \varepsilon\right)\log \rho,
\]

(4.13)

we can keep using Proposition 4.1 until we have only one box and its side length is \(\ell_2\). Then, we get

\[
\left(\frac{1}{2} - \varepsilon\right) \sum_{j=1}^{n} \int_{\Lambda_{\ell_2}} |W_j|^2 |\nabla j\phi_j(x_n)|^2 \, dx_n + 4\pi a \sum_{k=1}^{n} \int_{\Lambda_{\ell_2}} f(B_k) |\psi(x_n)|^2 \, dx_n
\geq
\]

\[
(1 - C\rho^2) |\log \rho| 4\pi a \int_{\Lambda_{\ell_2}} \frac{f_{s+1}(n)}{\ell^3_2} |\psi(x_n)|^2 \, dx_n.
\]

(4.14)

Note that we used an idea similar to (2.34) in order to convert the sum of kinetic energies over small cubic cells into the kinetic energy in a larger cubic cell.

By definition (4.5), \(K_{3h+1} > 2^{3h}\rho \ell^3_1 = \rho \ell^3_2\). Together with the definition of \(f\) in (2.20), it can be easily checked that \(f_{3h+1} \geq \bar{f}\). This proves the desired result. \(\Box\)

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 \(m_A, m_B = 0, 1, 2, \ldots, n\). Then,

\[
\min_{m_A + m_B = n} \frac{f_s(m_A) + f_s(m_B)}{|A_A|} \geq \frac{f_{s+1}(n)}{|A_A \cup A_B|}.
\]

(4.15)

We note that 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:

\[
\frac{4\pi a}{|\log \rho|} \sum_{j=1}^{n} \int_{(A_A \cup A_B)^n} |W_j|^2 |\nabla j\phi_j(x_n)|^2 \, dx_n + \int_{(A_A \cup A_B)^n} f_s(M(A)) + f_s(M(B)) |\psi(x_n)|^2 \, dx_n
\geq
\]

\[
(1 - C\rho^2) |\log \rho| 4\pi a \int_{(A_A \cup A_B)^n} f_{s+1}(n) |\psi(x_n)|^2 \, dx_n.
\]

(4.16)

Here, \(F_s\) is a function defined as follows:

**Definition 4.3.** For non-negative integers \(n_A, n_B\), define

\[
F_s(n_A, n_B, 0) := \frac{f_s(n_A)}{|A_A|} + \frac{f_s(n_B)}{|A_B|}.
\]

(4.17)

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

\[
P(y_i \in A_A) = P(y_i \in A_B) = \frac{1}{2}.
\]

(4.18)

Let \(m_A(k)\) and \(m_B(k)\) be the number of \(y_i\)'s in \(A_A\) and \(A_B\), respectively, i.e.,

\[
m_A(k) = \sum_{i=1}^{k} 1(y_i \in A_A), \quad m_B(k) = \sum_{i=1}^{k} 1(y_i \in A_B).
\]

(4.19)
Extend the definition of $F$ so that
\[
F_s(n_A, n_B, k) := \langle F_s(n_A + m_A(k), n_B + m_B(k), 0) \rangle_k,
\]
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 (4.10), 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,
\[
F_s(0, 0, n) \geq (1 - \rho) \frac{f_{s+1}(n)}{|\Lambda_A \cup \Lambda_B|},
\]
(4.21)

We note this lemma implies the second inequality of (4.10). 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 $x_n = (x_1, x_2, \cdots, x_n)$. Then, for any $\psi(x_n) \in L^2((\Lambda_A \cup \Lambda_B)^n)$ with $\phi_j = W_j^{-1} \psi$ for $j = 1, 2, \cdots, n$,
\[
\frac{(4\pi)^{-1}}{|\log \rho|} \sum_{j=1}^{n} \int_{(\Lambda_A \cup \Lambda_B)^n} |W_j|^2 |\nabla_j \phi_j(x_n)|^2 dx_n + \int_{(\Lambda_A \cup \Lambda_B)^n} F_s(\mathcal{M}(A), \mathcal{M}(B), 0) |\psi(x_n)|^2 dx_n \\
\geq (1 - C\rho^{\frac{1}{4}}) F_s(0, 0, n) \|\psi\|^2.
\]
(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 (4.10) from Lemma 4.3 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,
\[
\min_{m_A + m_B = n} f_s(m_A) + f_s(m_B) \geq 2 f_s(\frac{n}{2}).
\]
(4.23)

Thus, it suffices to prove
\[
2 f_s(\frac{n}{2}) \geq \frac{1}{2} f_{s+1}(n).
\]
(4.24)

(1) When $K_{s+1} + 2\sqrt{K_{s+1}} \leq n \leq 2K_s$, we have
\[
4 f_s(\frac{n}{2}) - f_{s+1}(n) = (n^2 - 2n) - [(2K_{s+1} - 1)n - K_{s+1}^2]
\]
\[
= (n - K_{s+1} - \frac{1}{2})^2 - K_{s+1} - \frac{1}{4}
\]
(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),
\[
2 \frac{d}{dn} \left( f_s(\frac{n}{2}) \right) = 2K_s - 1 \geq K_{s+1} - \frac{1}{2} = \frac{1}{2} \frac{d}{dn} \left( f_{s+1}(n) \right).
\]
(4.26)

Since we also have
\[
2 f_s(K_s) > \frac{1}{2} f_{s+1}(2K_s),
\]
(4.27)

We can see that, for any $n > 2K_s$,
\[
2 f_s(\frac{n}{2}) \geq \frac{1}{2} f_{s+1}(n).
\]
(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 $A_A \cup A_B$ so that, for $i = 1, 2, \cdots, n$,

$$P(y_i \in A_A) = P(y_i \in A_B) = \frac{1}{2}.$$  \hfill (4.29)

Let

$$m_A = \sum_{i=1}^{n} 1(y_i \in A_A), \quad m_B = \sum_{i=1}^{k} 1(y_i \in A_B),$$  \hfill (4.30)

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

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

To estimate the latter, we use the Chernoff bound (See Corollary 4.9 in [16].) for binomial distribution, which becomes, in this case,

$$P(m_A \geq n/2 + \zeta) \leq e^{-\frac{n\zeta^2}{4}}.$$  \hfill (4.33)

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

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

(4.33) implies

$$\langle m_A^2 - m_A - f_s(m_A) \rangle = \sum_{\zeta = K_s}^{\infty} P(m_A = \zeta) \cdot (\zeta - K_s)^2 \leq \sum_{\zeta = K_s}^{\infty} P(m_A \geq \zeta) \cdot (\zeta - K_s)^2 \leq \sum_{\zeta = K_s}^{\infty} e^{-\frac{2(\zeta - n/2)^2}{n}} (\zeta - K_s)^2 \leq \int_{K_s}^{\infty} e^{-\frac{2(\zeta - n/2)^2}{n}} (\zeta - K_s)^2 d\zeta \leq C n^2 e^{-\frac{n(2K_s - n/2)^2}{n}}.$$  \hfill (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

$$K_s - \frac{n}{2} \geq K_s - \frac{K_{s+1}}{2} - \sqrt{K_{s+1}} \gg \log \rho \frac{1}{2} \sqrt{K_{s+1}}.$$  \hfill (4.36)

Hence,

$$\frac{(K_s - n/2)^2}{n} \geq \frac{(K_s - n/2)^2}{2s + 2\rho^2 \rho_1^2} \gg |\log \rho| \frac{1}{2} \sqrt{K_{s+1}}.$$  \hfill (4.37)

which gives

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

Together with (4.35) we get

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

thus,

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

Therefore, from that

$$\langle f_s(m_A) + f_s(m_B) \rangle = 2\langle f_s(m_A) \rangle,$$  \hfill (4.41)
we obtain, when $n < K_{s+1} + 2\sqrt{K_{s+1}}$,
\begin{equation}
F_s(0,0,n) \geq \frac{1}{2} (1 - \rho) \frac{n^2 - n}{|A|} = (1 - \rho) \frac{n^2 - n}{|A \cup A|} \geq (1 - \rho) \frac{f_{s+1}(n)}{|A \cup A|},
\end{equation}
which was to be proved. \hfill \Box

### 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 $\frac{1}{2} \log \rho$.

**Lemma 4.6.** When $n < K_{s+1} + 2\sqrt{K_{s+1}}$, for any fixed $x_2, x_3, \ldots, x_N$ with $x_{n+1}, x_{n+2}, \ldots, x_N$ outside $(A \cup B)$, assume $n_A$ particles among $x_2, \ldots, x_n$ are in $A$ and $n_B$ particles in $B(n_A + n_B = n - 1)$. So, when $x_1$ is in $A$, the total energy is $F(n_A + 1, n_B, 0)$, otherwise it is $F(n_A, n_B, 1, 0)$. Let
\begin{equation}
T_1 = -W_1^{-1} \nabla_i W_1^2 \nabla_i W_1^{-1}
\end{equation}
defined on functions in $L^2(A \cup B)$ with the Neumann boundary conditions in $A \cup B$. Then, there exists a constant $C'$ such that, we have the following operator inequality:
\begin{equation}
\frac{(4\pi a)^{-1}}{|\log \rho|} T_1 + F(n_A + 1, n_B, 0) \cdot 1(x_1 \in A) + F(n_A, n_B, 1, 0) \cdot 1(x_1 \in B)
\geq F(n_A, n_B, 1) - C' \frac{(n_A - n_B)^2}{\ell^4 |\log \rho|^{-1}} - Cn\rho^{\frac{1}{4}}\ell^{-3}.
\end{equation}

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 $A \cup B$ 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 $-C'(n_A - n_B)^2\ell^{-2} |\log \rho|$. 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 $x_{k+1}, \ldots, x_n$ are in $A$ and $n_B$ particles in $B$ ($n_A + n_B = n - k$), we have
\begin{equation}
\frac{(4\pi a)^{-1}}{|\log \rho|} T_k + \left( F(n_A + 1, n_B, k - 1) - C'(k - 1) \frac{(n_A - n_B)^2}{\ell^4 |\log \rho|^{-1}} \right) \cdot 1(x_k \in A)
+ \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 B)
\geq F(n_A, n_B, k) - C' \frac{(n_A - n_B)^2}{\ell^4 |\log \rho|^{-1}} - Cn\rho^{\frac{1}{4}}\ell^{-3}.
\end{equation}

Here, $C$ does not depend on $k$ and $T_k$ is defined as
\begin{equation}
T_k = -W_k^{-1} \nabla_k W_k^2 \nabla_k W_k^{-1}
\end{equation}

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

**Proof of Lemma 4.5** Let
\begin{equation}
\mathcal{M}^{(k)}(A) = \sum_{i=k+1}^{n} 1(x_i \in A), \quad \mathcal{M}^{(k)}(B) = \sum_{i=k+1}^{n} 1(x_i \in B).
\end{equation}

Note that $\mathcal{M}(A) = \mathcal{M}^{(0)}(A)$ and $\mathcal{M}(B) = \mathcal{M}^{(0)}(B)$.
From Lemma 4.6 we get
\[
\frac{(4\pi a)^{-1}}{|\log \rho|} \int_{(A \cup A_B)^n} |W_1^2| \nabla_1 \phi_1(x_n)|^2 dx_n + \int_{(A \cup A_B)^n} F(M(A), M(B), 0) |\psi(x_n)|^2 dx_n
\] (4.48)
\[
\geq \int_{(A \cup A_B)^n} \left[ F(M^{(1)}(A), M^{(1)}(B), 1) - C r \frac{(M^{(1)}(A) - M^{(1)}(B))^2}{|\log \rho|^{-1}} - C n \rho \frac{\delta}{\ell^3} \right] |\psi(x_n)|^2 dx_n.
\]

As a next step, we apply Lemma 4.7 with the right hand side of the above equation. Then we get,
\[
\frac{(4\pi a)^{-1}}{|\log \rho|} \int_{(A \cup A_B)^n} |W_2^2| \nabla_2 \phi_2(x_n)|^2 dx_n
\]
\[
+ \int_{(A \cup A_B)^n} \left[ F(M^{(1)}(A), M^{(1)}(B), 1) - C r \frac{(M^{(1)}(A) - M^{(1)}(B))^2}{|\log \rho|^{-1}} \right] |\psi(x_n)|^2 dx_n
\] (4.49)
\[
\geq \int_{(A \cup A_B)^n} \left[ F(M^{(2)}(A), M^{(2)}(B), 2) - C r \frac{(M^{(2)}(A) - M^{(2)}(B))^2}{|\log \rho|^{-1}} - C n \rho \frac{\delta}{\ell^3} \right] |\psi(x_n)|^2 dx_n.
\]

Thus,
\[
\sum_{i=1,2} \frac{(4\pi a)^{-1}}{|\log \rho|} \int_{(A \cup A_B)^n} |W_i^2| \nabla_i \phi_i(x_n)|^2 dx_n + \int_{(A \cup A_B)^n} F(M(A), M(B), 0) |\psi(x_n)|^2 dx_n
\]
\[
\geq \int_{(A \cup A_B)^n} \left[ F(M^{(2)}(A), M^{(2)}(B), 2) - C r \frac{(M^{(2)}(A) - M^{(2)}(B))^2}{|\log \rho|^{-1}} - C n \rho \frac{\delta}{\ell^3} \right] |\psi(x_n)|^2 dx_n.
\]

We keep applying Lemma 4.7 Since $M^{(k)}(A) = M^{(k)}(B) = 0$, when $k = n$, we get
\[
\sum_{i=1}^n \frac{(4\pi a)^{-1}}{|\log \rho|} \int_{(A \cup A_B)^n} |W_i^2| \nabla_i \phi_i(x_n)|^2 dx_n + \int_{(A \cup A_B)^n} F(M(A), M(B), 0) |\psi(x_n)|^2 dx_n
\]
\[
\geq \int_{(A \cup A_B)^n} \left[ F(M^{(n)}(A), M^{(n)}(B), n) - C r \frac{(M^{(n)}(A) - M^{(n)}(B))^2}{|\log \rho|^{-1}} - C n^2 \rho \frac{\delta}{\ell^3} \right] |\psi(x_n)|^2 dx_n
\]
\[
= [F(0, 0, n) - C n^2 \rho \frac{\delta}{\ell^3}] \|\psi\|^2_2. (4.50)
\]

Since $F$ is convex, from Jensen’s inequality,
\[
F(0, 0, n) \geq F\left(\frac{n}{2}, \frac{n}{2}, 0\right) \geq C n^2 \ell^{-3}. (4.51)
\]

Thus, we have
\[
\frac{(4\pi a)^{-1}}{|\log \rho|} \sum_{i=1}^n \int_{(A \cup A_B)^n} |W_i^2| \nabla_i \phi_i(x_n)|^2 dx_n + \int_{(A \cup A_B)^n} F(M(A), M(B), 0) |\psi(x_n)|^2 dx_n
\]
\[
\geq (1 - C \rho \frac{\delta}{\ell^3}) F(0, 0, n) \|\psi(x_n)\|^2_2 dx_n, \tag{4.52}
\]
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
\[
g_k(x_k) = \begin{cases} 
F(n_A + 1, n_B, k - 1) - F(n_A, n_B, k - 1) & \text{if } x_k \in A_A \\
F(n_A, n_B + 1, k - 1) - F(n_A, n_B, k - 1) & \text{if } x_k \in A_B 
\end{cases} \tag{4.53}
\]
and
\[
M_k(x_k) = \begin{cases} 
(n_A + 1 - n_B)^2 - (n_A - n_B)^2 & \text{if } x_k \in A_A \\
(n_A - n_B - 1)^2 - (n_A - n_B)^2 & \text{if } x_k \in A_B 
\end{cases}. \tag{4.54}
\]
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.,
\[
\frac{(4\pi a)^{-1}}{|\log \rho|} T_k + g_k - C'(k-1) \frac{M_k}{\ell^4 |\log \rho|^{-1}} \geq F(n_A, n_B, k) - F(n_A, n_B, k-1) - \frac{C'}{\ell^4 |\log \rho|^{-1}} (n_A - n_B)^2 - Cn^\frac{3}{2}\ell^{-3}, \tag{4.56}
\]
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,
\[
g_k(x_k) = F(n_A + 1, n_B, k - 1) - F(n_A, n_B, k - 1) \tag{4.57}
\]
\[
= \langle F(n_A+1+ma(k-1), nB+mb(k-1), 0) - F(n_A+ma(k-1), nB+mb(k-1), 0) \rangle_{k-1}
\]
\[
= \frac{\langle f_s(n_A+1+ma(k-1)) - f_s(n_A+ma(k-1)) \rangle_{k-1}}{|A_A|} \tag{4.58}
\]
It can be easily checked that
\[
f_s(n_A+1+ma(k-1)) - f_s(n_A+ma(k-1)) \geq C(n_A+ma(k-1)).
\]
Thus,
\[
g_k(x_k) \geq Ck\ell^{-3}(n_A+ma(k-1))_{k-1} \geq Ck\ell^{-3}. \tag{4.59}
\]
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
\[
g_k \geq Ck\ell^{-3} \gg C \frac{k\rho}{\ell^4 |\log \rho|^{-1}} \geq \frac{kM_k}{\ell^4 |\log \rho|^{-1}}, \tag{4.60}
\]
which shows that the perturbation part of (4.55) is non-negative, i.e.,
\[
g_k - C'k\ell^{-3} \frac{M_k}{\ell^4 |\log \rho|^{-1}} \geq 0. \tag{4.61}
\]
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
\[
\frac{(4\pi a)^{-1}}{|\log \rho|} T_k + g_k - C'(k-1) \frac{M_k}{\ell^4 |\log \rho|^{-1}} \geq \langle g_k \rangle_{W_k} - C'(k-1) \frac{M_k}{\ell^4 |\log \rho|^{-1}} \tag{4.63}
\]
\[
- Ck^2 \frac{(M_k^2)_{W_k} - (M_k)_{W_k}^2}{\ell^3 |\log \rho|^{-2}} / (\ell^{-2} |\log \rho|^{-1})
\]
where \(\langle g_k \rangle_{W_k}\) denotes
\[
\langle g_k \rangle_{W_k} = \left( \int_{\Lambda_A \cup \Lambda_B} g_k(x_k)|W_k|^2 dx_k \right) / \left( \int_{\Lambda_A \cup \Lambda_B} |W_k|^2 dx_k \right) \tag{4.64}
\]
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 \(\langle g_k \rangle_{1_k}\) denote
\[
\langle g_k \rangle_{1_k} = \left( \int_{\Lambda_A \cup \Lambda_B} g_k(x_k) \cdot 1 \, dx_k \right) / \left( \int_{\Lambda_A \cup \Lambda_B} 1 \, dx_k \right)
\] (4.65)
and other expectations \(\langle \cdot \rangle_{1_k}\) are denoted similarly. It can be easily checked from the definitions that
\[
F(n_A, n_B, k - 1) + \langle g_k \rangle_{1_k} = F(n_A, n_B, k).
\] (4.66)
We have the following estimates for the terms in the right hand side of (4.63).

1. Using Lemma 5.3, we get
\[
\langle g_k \rangle_{W_k} - \langle g_k \rangle_{1_k} \leq Cn\rho\ell^2\epsilon^{-3}.
\] (4.67)
Together with (4.66), we have
\[
\langle g_k \rangle_{W_k} \geq F(n_A, n_B, k) - F(n_A, n_B, k - 1) - Cn\rho\ell^2\epsilon^{-3}.
\] (4.68)

2. Using Lemma 5.3, we get
\[
\left| \frac{M_k}{\ell^2 |\log \rho|^{-1}} \right| W_k - \left( \frac{M_k}{\ell^2 |\log \rho|^{-1}} \right)_{1_k} \leq Ck \frac{n\rho\ell^2}{\ell^2 |\log \rho|^{-1}} \ll n\rho\ell^2\epsilon^{-3}.
\] (4.69)
On the other hand, we simply follow the definitions to see
\[
\langle M_k \rangle_{1_k} = 1.
\] (4.70)
Hence, we get
\[
\left| \frac{M_k}{\ell^2 |\log \rho|^{-1}} \right| W_k \leq Cn\rho\ell^2\epsilon^{-3}.
\] (4.71)

3. Using Lemma 5.3, we get
\[
\left| \langle g_k^2 \rangle_{W_k} - \langle g_k^2 \rangle_{1_k} \right| \leq Cn^2 \rho \ell^2 \epsilon^{-6},
\] (4.72)
\[
\left| \langle g_k^2 \rangle_{W_k} - \langle g_k^2 \rangle_{1_k} \right| \leq Cn^2 \rho \ell^2 \epsilon^{-6}.
\] (4.73)
Thus, since \(n \leq C\rho\ell^3 \ll \ell |\log \rho|^{-1}\),
\[
\left| \frac{\langle g_k^2 \rangle_{W_k} - \langle g_k^2 \rangle_{1_k}}{\ell^{-2} |\log \rho|^{-1}} \right| \leq \left| \frac{\langle g_k^2 \rangle_{1_k} - \langle g_k^2 \rangle_{1_k}}{\ell^{-2} |\log \rho|^{-1}} \right| + Cn\rho\ell^2\epsilon^{-3}.
\] (4.74)
To estimate \(\langle g_k^2 \rangle_{1_k} - \langle g_k^2 \rangle_{1_k}\), we again follow the definition and get
\[
\langle g_k^2 \rangle_{1_k} - \langle g_k^2 \rangle_{1_k} = \left[ F(n_A + 1, n_B + 1, k - 1) - F(n_A, n_B + 1, k - 1) \right]^2.
\] (4.75)
Suppose that we have \((k - 1)\) particles randomized, namely \(y_1, y_2, \ldots, y_{k-1}\) as in the definition of \(F\), (4.17) - (4.20). Then,
\[
\left[ F(n_A + 1, n_B, k - 1) - F(n_A, n_B, k - 1) \right]^2
\] (4.76)
\[
= \left[ F(n_A + 1 + m_A(k - 1), n_B + m_B(k - 1), 0) - F(n_A + m_A(k - 1), n_B + 1 + m_B(k - 1), 0) \right]_{k-1}^2
\] (4.77)
\[
\leq \left[ F(n_A + 1 + m_A(k - 1), n_B + m_B(k - 1), 0) - F(n_A + 1 + m_A(k - 1), n_B + 1 + m_B(k - 1), 0) \right]_{k-1}^2
\] (4.78)
It can be easily checked that, for any non-negative integers \(m_1\) and \(m_2\),
\[
\left| F(m_1 + 1, m_2, 0) - F(m_1, m_2 + 1, 0) \right|^2 \leq \frac{4(m_1 - m_2)^2}{|\Lambda_A|}.
\] (4.77)
Thus,
\[
\left[ F(n_A + 1, n_B + 1, k - 1) - F(n_A, n_B + 1, k - 1) \right]^2
\] (4.76)
\[
\leq 4|\Lambda_A|^{-2} \left[ \left( (n_A + m_A(k - 1)) - (n_B + m_B(k - 1)) \right)^2 \right]_{k-1}.
\] (4.78)
From the definition,
\[ \langle [ (n_A + m_A(k-1)) - (n_B + m_B(k-1))^2 ] \rangle_{k-1} = (n_A - n_B)^2 + (k-1). \] (4.79)

Thus, from (4.73), (4.75), (4.78), and (4.79), we obtain
\[ \frac{\langle g_k^2 \rangle_w - \langle g_k^2 \rangle_{w_k}}{\ell^{-2} |\log \rho|^{-1}} \leq |\Lambda_1|^{-2} (n_A - n_B)^2 + k \leq C (n_A - n_B)^2 + Cn \rho^{1/2} \ell^{-3}. \] (4.80)

(4) Using Lemma 5.3 from that \( n \leq C \rho^{3} \ll \ell |\log \rho|^{-1} \) we get
\[ k^2 \frac{\langle M_k^2 \rangle_{w_k} - \langle M_k^2 \rangle_{w_k}}{\ell^6 |\log \rho|^{-3}} \leq \frac{C n^4}{\ell^6 |\log \rho|^{-3}} \ll n \rho^{1/2} \ell^{-3}. \] (4.81)

From the definition of \( M_k \), it can be easily checked that
\[ \langle M_k^2 \rangle_{1_k} - \langle M_k \rangle_{1_k}^2 = 4(n_A - n_B)^2. \] (4.82)

Thus,
\[ k^2 \frac{\langle M_k^2 \rangle_{1_k} - \langle M_k \rangle_{1_k}^2}{\ell^6 |\log \rho|^{-3}} = k^2 \frac{4(n_A - n_B)^2}{\ell^6 |\log \rho|^{-3}} \leq \frac{C n^4}{\ell^6 |\log \rho|^{-3}} \ll n \rho^{1/2} \ell^{-3}. \] (4.83)

Here, the last inequality follows from
\[ n^3 \ell^{-3} |\log \rho|^3 \leq C \rho^{3} \ell^6 |\log \rho|^3 \ll \rho^{1/2}. \] (4.84)

Hence,
\[ k^2 \frac{\langle M_k^2 \rangle_{w_k} - \langle M_k \rangle_{w_k}^2}{\ell^6 |\log \rho|^{-3}} \leq n \rho^{1/2} \ell^{-3}. \] (4.85)

Applying estimates (4.78), (4.79), (4.80), and (4.83) to (4.85), we obtain (4.80). 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 \( \phi \) be the lowest Neumann eigenvalue and eigenfunction on the ball of radius \( \kappa \), i.e.,
\[ (-\Delta + \frac{1}{2} V) \phi = e_0 \phi \] (5.1)
with the boundary condition
\[ \phi(x) = 1 \text{ and } \partial \phi(x) = 0 \text{ if } |x| = \kappa. \] (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
\[ c_0 \leq \phi(x) \leq 1, \quad 1 - \phi(x) \leq \frac{c_1}{|x|}. \] (5.3)

Moreover, we have
\[ \frac{3a}{\kappa^3} \leq e_0 \leq \frac{3a}{\kappa^3} (1 + \frac{C}{\kappa}). \] (5.4)

Proof. For a proof of the first part of the lemma,
\[ c_0 \leq \phi(x) \leq 1, \quad 1 - \phi(x) \leq \frac{c_1}{|x|}, \] (5.5)
see Lemma A.1 in [3]. To prove the second part, we extend the definition of \( \phi \) so that \( \phi(x) = 1 \) if \( |x| > \kappa \). It is well known that
\[ \int_{\mathbb{R}^3} \phi(x) \left( -\Delta + \frac{1}{2} V(x) \right) \phi(x) dx \geq 4\pi a. \] (5.6)
Thus, $\phi(x) \leq 1$ implies that
\begin{equation}
    c_0 \geq 4\pi a \left( \int_{|x| \leq \kappa} |\phi(x)|^2 dx \right)^{-1} \geq \frac{3a}{\kappa^3}.
\end{equation}
Upper bound for $c_0$ is also proved in Lemma A.1 in [3].

**Lemma 5.2.** Define $W_j$ as in [4.14] and $x = (x_1, x_2, \ldots, x_N)$. Suppose that $x_j$ is in $\Lambda_\ell$, a box of side length $\ell$ and $x_1, \ldots, x_j, \ldots, x_N$ are fixed.

1. There exists a constant $c_0 < 1$ such that
\begin{equation}
    W_j(x) \geq 1 - c_0.
\end{equation}

2. Let $\ell \gg \ell_0$. If $\Lambda_\ell$ contains $n$ particles including $x_j$, then
\begin{equation}
    \ell^3 (1 - Cn \frac{\ell^2}{\ell^3} - C\ell^{-1}) \leq \int_{\Lambda_\ell} |W_j|^2 dx_j \leq \ell^3
\end{equation}

**Proof.** From Lemma 5.1, we can see that there exists a constant $c_0$ that does not depend on $\kappa$ such that $\tau(\kappa, x_i - x_j) \leq c_0$ whenever $\ell - 1 \leq \kappa \leq \ell_0$. It is clear that
\begin{equation}
    F_{ij}(x) \tau(t_0, x_i - x_j) > 0 \quad \text{or} \quad (1 - F_{ij}(x))G_{ij}(x)\tau(t_{ij}, x_i - x_j) > 0
\end{equation}
implies that $x_i$ is the nearest particle of $x_j$. Thus, there exists an index $k$ such that
\begin{equation}
    W_j(x) = 1 - \left[ F_{kj}(x) \tau(t_0, x_k - x_j) + (1 - F_{kj}(x))G_{kj}(x)\tau(t_{kj}, x_k - x_j) \right],
\end{equation}
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 $t_{kj} > \ell_0$,
\begin{equation}
    \int_{|x_k - x_j| < \ell_0} |W_j|^2 dx_j \geq \int_{|x_k - x_j| < \ell_0} \left( 1 - \frac{C}{|x_k - x_j|} \right)^2 dx_j \geq \int_{|x_k - x_j| < \ell_0} dx_j - C\ell_0^2,
\end{equation}
and, when $t_{kj} \leq \ell_0$,
\begin{equation}
    \int_{|x_k - x_j| < t_{kj}} |W_j|^2 dx_j \geq \int_{|x_k - x_j| < t_{kj}} \left( 1 - \frac{C}{|x_k - x_j|} \right)^2 dx_j \geq \int_{|x_k - x_j| < t_{kj}} dx_j - C\ell_{ij}^2 \geq \frac{4}{3} \pi r_k^3 - C\ell_{ij}^2.
\end{equation}

Now consider the other case where $x_k \notin \Lambda_\ell$. Note that $d(x_k, \Lambda_\ell) \leq \ell_0$ in this case, where $d(x_k, \Lambda_\ell)$ denotes the distance from $x_k$ to the box $\Lambda_\ell$. When $t_{kj} > \ell_0$, let $B_k$ be a ball of radius $\ell_0$ centered at $x_k$. $B_k \cap \partial \Lambda_\ell$, the intersection of $B_k$ and the boundary of $\Lambda_\ell$, is a disk, and we let $r_k$ be the radius of the disk. Since
\begin{equation}
    \int_{B_k \cap \Lambda_{\ell}} \frac{1}{|x_j - x_k|} dx_j \leq \int_{\sqrt{\ell_0^2 - r_k^2} \leq |x_j - x_k| \leq \ell_0} \frac{1}{|x_j - x_k|} dx_j = 2\pi r_k^2,
\end{equation}
we can see that
\begin{equation}
    \int_{B_k \cap \Lambda_{\ell}} |W_j|^2 dx_j \geq \int_{B_k \cap \Lambda_{\ell}} \left( 1 - \frac{C}{|x_j - x_k|} \right)^2 dx_j \geq \int_{B_k \cap \Lambda_{\ell}} dx_j - Cr_k^2.
\end{equation}
We can get the same estimate when $t_{kj} \leq \ell_0$, 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
\begin{equation}
    \sum_{k : x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \leq \ell_0} \int_{B_k \cap \Lambda_{\ell}} |W_j|^2 dx_j \geq \sum_{k : x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \leq \ell_0} \int_{B_k \cap \Lambda_{\ell}} dx_j - C\sum_{k : x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \leq \ell_0} r_k^2.
\end{equation}
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,
\begin{equation}
    \sum_{k : x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \leq \ell_0} \int_{B_k \cap \Lambda_{\ell}} |W_j|^2 dx_j \geq \sum_{k : x_k \notin \Lambda_\ell, d(x_k, \Lambda_\ell) \leq \ell_0} \int_{B_k \cap \Lambda_{\ell}} dx_j - C\ell^2.
\end{equation}
Altogether, we get
\[
\int_{\Lambda} |W_j|^2 dx_j \geq \int_{\Lambda} dx_j - C n \ell_0^2 - C \ell^2 = \ell^3 \left(1 - C n \frac{\ell_0^2}{\ell^3} - C \ell^{-1}\right).
\] (5.18)

Finally, \(W_j \leq 1\) implies that
\[
\int_{\Lambda} |W_j|^2 dx_j \leq \ell^3.
\] (5.19)

This proves the second part of the lemma. \(\square\)

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

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

let

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

\[
\langle h \rangle_{1_k} = \int_{\Lambda_A \cup \Lambda_B} h(x_j) \cdot 1 dx_j / \int_{\Lambda_A \cup \Lambda_B} 1 dx_j.
\] (5.22)

Then,

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

**Proof.** Let

\[
p_A = \left(\int_{\Lambda} |W_j|^2 dx_j \right) / \left(\int_{\Lambda_A \cup \Lambda_B} |W_j|^2 dx_j \right),
\] (5.24)

\[
p_B = \left(\int_{\Lambda_B} |W_j|^2 dx_j \right) / \left(\int_{\Lambda_A \cup \Lambda_B} |W_j|^2 dx_j \right).
\] (5.25)

Then,

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

From Lemma 5.2,

\[
p_A = \frac{1}{2} + O\left(\frac{n \ell_0^2}{\ell^3}\right) + O(\ell^{-1}) = \frac{1}{2} + O(\rho \ell_0^2) + O(\ell^{-1}),
\] (5.27)

and the same estimate holds for \(p_B\). Since

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

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

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

which was to be proved. \(\square\)

**Acknowledgment**

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


(Ji Oon Lee) DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138, USA
E-mail address: jion@math.harvard.edu

(Jun Yin) DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138, USA
E-mail address: jyin@math.harvard.edu