Accuracy of the Hartree-Fock Approximation for the Hubbard Model

V. Bach, J. Poelchau

February 23, 1996

Abstract

A lower bound $0 \geq \big(E_{gs}(n) - E_{hf}(n)\big)/|\Lambda| \geq -\text{const}\,\big[n^{2/3}\,U^{4/3}(|\ln U| + 1) + Un^{1/2}|\Lambda|^{-1/2d}(|\ln(|\Lambda|^{-1/2})| + 1)\big]$ to the difference of the ground state- and the Hartree-Fock energy of the Hubbard model is derived. $|\Lambda|$ is the lattice size, U is the coupling parameter, and n is the electron density per site. This estimate holds for all dimensions $d \geq 2$ and all densities. Thus the Hartree-Fock approximation becomes exact (even beyond terms of order U) for small U and large $|\Lambda|$.

1 Introduction

The Hubbard model describes electrons on a finite hypercubic lattice $\Lambda = [-L+1, L]^d \cap \mathbb{Z}^d$. We assume periodic boundary conditions, i.e., $\Lambda = (\mathbb{Z}/2L\mathbb{Z})^d$. The one-particle Hilbert space consists of the spinor-valued functions on Λ: $\mathcal{H} = \{f : \Lambda \times \{\uparrow,\downarrow\} \to \mathbb{C}\} = \mathbb{C}^{|\Lambda|} \otimes \mathbb{C}^2$. The Hamiltonian is usually given in second quantized form

$$H := \sum_{\substack{(x,y)\in\Lambda^2 \\ \sigma\in\{\uparrow,\downarrow\}}} t_{xy}\, c^\dagger_{x,\sigma} c_{y,\sigma} \; + \; U \sum_{x\in\Lambda} n_{x,\uparrow} n_{x,\downarrow} \qquad (U > 0)\ . \tag{1}$$

It is defined on the Fermion Fockspace $\mathcal{F}(\mathcal{H})$ over $\mathcal{H}$. We denote the vacuum by Ω. Vectors of the form $\prod_{n=1}^N c^\dagger_{x_n,\sigma_n}\Omega$ generate $\mathcal{F}(\mathcal{H})$, where $c^\dagger_{x,\sigma}$ are the usual Fermion creation operators. The canonical ONB $\{\delta_{x,\sigma} \mid \delta_{x,\sigma}(y,\tau) = \delta_{xy}\delta_{\sigma\tau},\ x \in \Lambda,\ \sigma = \uparrow,\downarrow\} \subset \mathcal{H}$ may be identified with the vectors $c^\dagger_{x,\sigma}\Omega \in \mathcal{F}(\mathcal{H})$. We assume the hopping matrix t to be translational invariant: $t_{xy} = t(x-y)$. The operator $n_{x,\sigma} = c^\dagger_{x,\sigma}c_{x,\sigma}$ registers an electron with spin σ at the site x, so the second term in (1) describes the Coulomb repulsion of two electrons located at the same site. It is believed that the Hubbard model describes important features of condensed matter systems, namely the competition between the kinetic energy and the Coulomb interaction, in the simplest way. One is interested in the basic properties (spectral data, magnetic ordering, charge transport) of its ground- and Gibbs states. Of special interest for the theory of high temperature superconductors, in which magnetism and superconductivity are intimately related, is the magnetic structure of the these states if the particle number is $N = |\Lambda|$ (half filling) and away from half filling ($N \neq |\Lambda|$). In spite of the simplicity of the Hubbard model, only little is known about its ground state for dimension greater than one. For this reason, various approximations depending on the coupling constant and the density $n = N/|\Lambda|$ have become customary in the physics literature. Their mathematical justification is far from obvious.

In the present paper we show that the Hartree-Fock (HF) approximation (E_{hf}) of the ground state energy (E_{gs}) becomes exact in the weak coupling limit for all particle numbers and all dimensions $d \geq 2$:

Theorem

Let t be a translational invariant hopping matrix with finite support, such that $\hat{t} : [-\pi, \pi]^d \to \mathbb{C}$ is a Morse function, and let Λ be a lattice such that $\operatorname{supp} t \subset \Lambda$. *Let $n := N/|\Lambda|$ be the density. If the dimension of the lattice Λ is $d = 2$ then*

$$0 \geq \frac{E_{gs}(n) - E_{hf}(n)}{|\Lambda|} \geq -\text{const} \left(n^{2/3} U^{4/3} (|\ln U| + 1) \;+\; n^{1/2} U |\Lambda|^{-1/4} (|\ln(|\Lambda|^{-1/2})| + 1) \right) \quad (2)$$

holds true, and if $d \geq 3$

$$0 \geq \frac{E_{gs}(n) - E_{hf}(n)}{|\Lambda|} \geq -\text{const} \left(n^{2/3} U^{4/3} \;+\; n^{1/2} U |\Lambda|^{-1/2d} \right) . \quad (3)$$

Note that with our condition of finite support of the hopping matrix the Fourier transform $\hat{t}$ is defined independently of the lattice Λ (provided it is large enough and contains $\operatorname{supp} t$) and may thus be viewed as a function on the continuous Brillouin zone $[-\pi, \pi]^d$. We point out that the dependence of the errors in (2) and (3) on the lattice size $|\Lambda|$ is explicit. Thus a limit $U \to 0$ is meaningful even for a sequence of finite lattices, if we require $|\Lambda| \gg U^{-2d/3}$. Writing the Hubbard Hamiltonian as $H =: T + UW$, where T denotes the kinetic term and UW denotes the interaction term, it is easily seen that

$$E_{hf}(n) = \langle \Phi_{hf} | T \Phi_{hf} \rangle + \langle \Phi_{hf} | UW \Phi_{hf} \rangle = O(|\Lambda|) + O(U|\Lambda|) \ , \quad (4)$$

where Φ_{hf} is the HF-ground state (see Section 2). In the HF-approximation the exchange correlations are taken into account. The theorem states that for small U and large $|\Lambda|$ the higher correlations contribute the small error $\text{const}\, n^{2/3} U^{4/3} |\ln U|$ at most. To prove the theorem above we use the correlation inequality from [4], which is a generalization of the one in [1] that was developed for a proof of an analogous statement for atoms and molecules. This type of correlation inequality allows for an estimate of electron pair correlations in terms of single particle quantities. An important condition for the applicability of the correlation inequalities is the representability of the interaction term as a sum (or integral) of projections (see (8)), which implies the positivity of the potential and its Fourier transform. Although the correlation estimate itself does not depend on the spatial dimension d, the lower bound we obtain from its application does. In fact, in the case of the Hubbard model the density of states (DOS) of the free ($U = 0$) model enters (see Section 3). The logarithmic terms in (2) arise from the Van Hove singularities of the DOS. We remark that our estimates are uniform in the density n in contrast to the one dimensional case where one would have to exclude densities corresponding to critical values of the dispersion relation (see [2], Thm. 3).

2 Hartree-Fock Energy and Correlation Inequalities

To formulate and understand the problem of the Hartree-Fock approximation of the ground state energy we introduce some notation, some of which is standard and some is not. For a fixed density $n = N/|\Lambda|$ (we shall assume $n = N/|\Lambda|$ without further notice) the **ground state energy** is defined by

$$E_{gs}(n) := \inf\{\langle \Psi \mid H\Psi\rangle \mid \ \Psi \in \mathcal{F}(\mathcal{H}) \ , \ \|\Psi\| = 1 \ , \ \mathbf{n}\Psi = n\Psi\} \ , \tag{5}$$

where $\mathbf{n} = \mathbf{N}/|\Lambda|$ is the density operator ($\mathbf{N} = \sum_{x\in\Lambda,\sigma} c^\dagger_{x,\sigma}c_{x,\sigma}$ is the number operator). A normalized vector $\Psi_{gs} \in \mathcal{F}(\mathcal{H})$ with the properties $\mathbf{n}\Psi_{gs} = n\Psi_{gs}$ and $\langle \Psi_{gs} \mid H\Psi_{gs}\rangle = E_{gs}(n)$ is called a **ground state** of H. The **Hartree-Fock** (HF) **energy** is defined by

$$E_{hf}(n) := \inf\{ \langle \Phi \mid H\Phi\rangle \mid \Phi \in \mathcal{SD}_{n|\Lambda|}\} \ , \tag{6}$$

where $\mathcal{SD}_{n|\Lambda|} = \left\{ \Phi \,\middle|\, \exists f_1,\ldots,f_{n|\Lambda|} \in \mathcal{H},\ \langle f_i|f_j\rangle = \delta_{ij},\ \Phi = \prod_{j=1}^{n|\Lambda|} c^\dagger(f_j)\Omega \right\}$ is the set of $n|\Lambda|$-particle Slater determinants. A vector $\Phi_{hf} \in \mathcal{SD}_{n|\Lambda|}$ with the property $E_{hf}(n) = \langle \Phi_{hf}|H\Phi_{hf}\rangle$ is called a **Hartree-Fock ground state** of H.

The two main goals in an approximate theory like HF-theory are clearly the following:

- Justify the application of the approximate theory.
- Calculate the HF-energy and the properties of the HF-ground-states. Here we have the following situation:
  - At half filling ($n = 1$) it has been shown [3] that Φ_{hf} is an antiferromagnetic insulator with AF-Nèel order. This result holds for general bipartite lattices. In the translational invariant case, which we are considering here, the HF-energy is accessible to explicit computation, resulting from a BCS equation.
  - If $n \neq 1$, few is known about the magnetic structure of the HF-ground-states but it is believed that the antiferromagnetism of the HF-ground-states for $n = 1$ is unstable against the injection of electrons or holes (see [6]).

In this paper we are concerned with the first item, and we consider the energy difference

$$\triangle E(n) := E_{gs}(n) - E_{hf}(n)\,. \tag{7}$$

Clearly $\triangle E(n) < 0$. If we want to obtain a lower bound for $\triangle E(n)$ we are confronted with the problem of minimizing the energy 'functional' $\langle\, (\cdot) \mid H(\cdot)\,\rangle$ on two different sets (the normalized $n|\Lambda|$-particle vectors and the $n|\Lambda|$-particle Slater determinants). Rewriting the two variational problems in terms of reduced density matrices clarifies their structural difference, and we can easily derive an explicit expression for $\triangle E(n)$, which we estimate with the help of the correlation inequalities. To express $\triangle E(n)$ explicitly we write H in standard form:

$$H \ = \ \sum_{xy,\sigma\tau} h_{x\sigma,y\tau}\, c^\dagger_{x,\sigma}c_{y,\tau} \ + \ \frac{1}{2}\sum_{\substack{xy,\tilde{x}\tilde{y}\\ \sigma\tau,\tilde{\sigma}\tilde{\tau}}} V_{x\sigma,y\tau\,;\,\tilde{x}\tilde{\sigma},\tilde{y}\tilde{\tau}}\ c^\dagger_{y,\tau}c^\dagger_{x,\sigma}c_{\tilde{x},\tilde{\sigma}}c_{\tilde{y},\tilde{\tau}} \ = \ d\Gamma(h) + d\Gamma(V) \ \ ,$$

where $d\Gamma(\cdot)$ denotes the second quantization of the operators $h : \mathcal{H} \to \mathcal{H}$ and $V : \mathcal{H}\otimes\mathcal{H} \to \mathcal{H}\otimes\mathcal{H}$. The kinetic term h is determined by the hopping matrix: $h = t \otimes \mathbf{1}_{\mathbb{C}^2}$. The repulsion term V has the required decomposition property mentioned above:

$$V = U\sum_{x\in\Lambda} X_x \otimes X_x\ , \quad \text{where} \quad X_x = \sum_{\sigma} |\delta_{x\sigma}\rangle\langle\delta_{x\sigma}|\ . \tag{8}$$

Note that X_x are orthogonal projections obeying

$$\sum_{x\in\Lambda} X_x = \mathbf{1} \tag{9}$$

To a normalized $\Psi \in \mathcal{F}$ we associate an operator γ_Ψ on $\mathcal{H}$ and an operator Γ_Ψ on $\mathcal{H}\otimes\mathcal{H}$. They are given by

$$\begin{aligned}
\gamma_\Psi &:= \sum_{k;m} \langle \Psi \mid c^\dagger(\varphi_k) c(\varphi_m) \Psi\rangle \;\; |\varphi_m\rangle\langle\varphi_k| \;\; , \\
\Gamma_\Psi &:= \sum_{k,l;m,n} \langle \Psi \mid c^\dagger(\varphi_l) c^\dagger(\varphi_k) c(\varphi_m) c(\varphi_n) \Psi\rangle \;\; |\varphi_m\otimes\varphi_n\rangle\langle\varphi_k\otimes\varphi_l| \;\; ,
\end{aligned}$$

where $\{\varphi_k\}_{k\in\{1,\dots,2|\Lambda|\}}$ is some ONB in $\mathcal{H}$. γ_Ψ is called the **one-particle density matrix** (1-pdm) of Ψ and Γ_Ψ is called the **two-particle density matrix** (2-pdm) of Ψ. Let us denote by Tr_1 and Tr_2 the traces over $\mathcal{H}$ and $\mathcal{H}\otimes\mathcal{H}$, respectively. We recall some general properties of the 1-and 2-pdm's. Both are positive operators: $0 \le \gamma_\Psi = \gamma_\Psi^\dagger \le 1$, $0 \le \Gamma_\Psi = \Gamma_\Psi^\dagger \le \mathrm{Tr}_2[\Gamma_\Psi]$, and it holds $\mathrm{Tr}_1[\gamma_\Psi] = \langle\Psi \mid \mathbf{N}\Psi\rangle$. If $\Phi \in SD_{n|\Lambda|}$ is a given Slater determinant then $\gamma_\Phi = \gamma_\Phi^2$ is an orthogonal projection of rank $\mathrm{Tr}_1[\gamma_\Phi] = n|\Lambda|$. The 2-pdm corresponding to Φ is given by $\Gamma_\Phi = (\mathbf{1} - \mathrm{Ex})(\gamma_\Phi\otimes\gamma_\Phi)$, where $\mathrm{Ex} : f\otimes g \mapsto g\otimes f$ is the **exchange operator** on $\mathcal{H}\otimes\mathcal{H}$. The proof of these statements is straightforward.

Let γ_{gs} and Γ_{gs} be the 1- and 2-pdm's corresponding to a ground state Ψ_{gs} of H for $U > 0$. We may rewrite the ground state energy in (5) as

$$E_{gs}(n) = \mathrm{Tr}_1[h\gamma_{gs}] + \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}] \; .$$

Additionally the HF energy may be expressed as

$$\begin{aligned}
E_{hf}(n) &= \inf\left\{ \mathrm{Tr}_1[h\gamma_\Phi] + \frac{1}{2}\mathrm{Tr}_2[V\Gamma_\Phi] \;\middle|\; \Phi \in \mathcal{SD}_{n|\Lambda|} \right\} \\
&= \inf\left\{ \mathrm{Tr}_1[h\gamma_\Phi] + \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1}-\mathrm{Ex})(\gamma_\Phi\otimes\gamma_\Phi)] \;=:\; \mathcal{E}_{hf}(\gamma_\Phi) \;\middle|\; \Phi \in \mathcal{SD}_{n|\Lambda|} \right\} \; .
\end{aligned}$$

$\mathcal{E}_{hf}$ is called the HF energy functional. Now we have worked out the set-up to estimate $\triangle E(n)$ explicitly. Let γ_0 be the 1-pdm corresponding to the ground state Φ_0 of the free ($U = 0$) model, obviously Φ_0 is a Slater determinant. Intuitively γ_{gs} 'converges' to γ_0 if U tends to zero (γ_{gs} is the 1-pdm corresponding to the ground state for finite coupling $U > 0$). This we are going to quantify in a weak sense. The first step is the following

Lemma 1

$$\triangle E(n) \ge \frac{U}{2}\sum_{x\in\Lambda} \mathrm{Tr}_2\Big[(X_x\otimes X_x)\{\Gamma_{gs} - (\mathbf{1} - \mathrm{Ex})(\gamma_0\otimes\gamma_0)\}\Big] \tag{10}$$

Proof: We will use two simple facts to estimate $\triangle E(n)$:

$$E_{hf}(n) \;\le\; \mathrm{Tr}_1[h\gamma_0] + \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1}-\mathrm{Ex})(\gamma_0\otimes\gamma_0)] \; , \text{ and} \tag{11}$$

$$\mathrm{Tr}_1[h\gamma_{gs}] - \mathrm{Tr}_1[h\gamma_0] \;\ge\; 0 \; . \tag{12}$$

Indeed, (11) follows from inserting $\Phi_0 \in \mathcal{SD}_{n|\Lambda|}$ into (6) as a trial state, and (12) is reexpressing the fact that Φ_0 is a ground state for $U = 0$. Inequalities (11) and (12) imply that

$$\begin{aligned}
0 \ \geq \triangle E(n) \ &\geq \ E_{gs}(n) - \mathrm{Tr}_1[h\gamma_0] - \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1} - \mathrm{Ex})(\gamma_0 \otimes \gamma_0)] \\
&= \ \mathrm{Tr}_1[h\gamma_{gs}] - \mathrm{Tr}_1[h\gamma_0] + \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}] - \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1} - \mathrm{Ex})(\gamma_0 \otimes \gamma_0)] \\
&\geq \ \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}] - \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1} - \mathrm{Ex})(\gamma_0 \otimes \gamma_0)] \\
&= \ \frac{U}{2}\sum_{x\in\Lambda} \mathrm{Tr}_2\Big[(X_x \otimes X_x)\{\Gamma_{gs} - (\mathbf{1} - \mathrm{Ex})(\gamma_0 \otimes \gamma_0)\}\Big]
\end{aligned}$$

In the last step have used (8). ■

We want to remark that we do not need Lieb's variational principle [5, 1] to achieve (10), nor elsewhere in the present application. To estimate the right hand side of (10) we use a result by Graf and Solovej (see [4] Theorem 3 and Corrolary 5).

Lemma 2 [4] *Let X and P be bounded operators on a Hilbert space $\mathcal{H}$ satisfying $X = X^2 = X^\dagger$ and $0 \leq P \leq 1$, and let $\xi \geq 0$, then*

$$\begin{aligned}
&\frac{1}{2}\ \mathrm{Tr}_2[(X \otimes X)\Gamma_\Psi] - \xi\ \mathrm{Tr}_1[X\gamma_\Psi] \ + \ \frac{\xi^2}{2} \ + \ \frac{1}{2}\ \mathrm{Tr}_1[PXPX] \ \geq \\
&\qquad -\mathrm{const}\,\Big(\xi + \ \mathrm{Tr}_1[X\gamma_\Psi] + \ \mathrm{Tr}_1[XP]\Big)\min\Big\{1\,,\ \sqrt{\ \mathrm{Tr}_1[X(\mathbf{1} - P)\gamma_\Psi(\mathbf{1} - P)]}\ \Big\} \qquad (13)
\end{aligned}$$

for any normalized $\Psi \in \mathcal{F}(\mathcal{H})$.

For our purpose we set $\xi = \mathrm{Tr}_1[X\gamma_0]$ and $P = \gamma_0$ which makes it possible to estimate $\mathrm{Tr}_2\Big[(X_x \otimes X_x)\{\Gamma_{gs} - (\mathbf{1} - \mathrm{Ex})(\gamma_0 \otimes \gamma_0)\}\Big]$.

Lemma 3

$$\frac{\triangle E(n)}{|\Lambda|} \geq -\mathrm{const}\ Un^{1/2}\left(\frac{\mathrm{Tr}_1[(\mathbf{1} - \gamma_{gs})\gamma_0]}{|\Lambda|}\right)^{1/2} \ . \qquad (14)$$

Lemma 3 is an important step in our estimate. It is already visible that we have achieved a lower bound which is better than linear in U, because one expects the 'overlap' $\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_0]$ to vanish in some order in U as $U \to 0$.

Proof of Lemma 3:

We rewrite (10) as

$$\begin{aligned}
\frac{1}{U}\triangle E(n) \ &\geq \ \frac{1}{2}\left(\sum_{x\in\Lambda} \mathrm{Tr}_2[(X_x \otimes X_x)\Gamma_{gs}] - 2\big(\mathrm{Tr}_1[X_x\gamma_0]\big)^2 + \mathrm{Tr}_1[X_x\gamma_0 X_x\gamma_0]\right) \\
&= \ \sum_{x\in\Lambda}(A_x + B_x)\ , \qquad (15)
\end{aligned}$$

where

$$\begin{aligned}
A_x \ &:= \ \frac{1}{2}\Big(\mathrm{Tr}_2[(X_x \otimes X_x)\Gamma_{gs}] - 2\mathrm{Tr}_1[X_x\gamma_0]\mathrm{Tr}_1[X_x\gamma_{gs}] + \big(\mathrm{Tr}_1[X_x\gamma_0)]\big)^2\Big)\ , \text{ and} \qquad (16) \\
B_x \ &:= \ \mathrm{Tr}_1[X_x\gamma_0]\mathrm{Tr}_1[X_x(\gamma_{gs} - \gamma_0)]\ . \qquad (17)
\end{aligned}$$

Now we apply Lemma 2 with $\xi = \mathrm{Tr}_1[X_x \gamma_0]$ and $P = \gamma_0$. This yields

$$A_x \geq -\mathrm{const}\, \mathrm{Tr}_1[X_x(\gamma_{gs} + 2\gamma_0)] \sqrt{\mathrm{Tr}_1[X_x(\mathbf{1}-\gamma_0)\gamma_{gs}(\mathbf{1}-\gamma_0)]} \ .$$

Using Schwarz's inequality we can eliminate the projections X_x.

$$\begin{aligned}
\sum_{x\in\Lambda} A_x &\geq -\mathrm{const} \left(\sum_{x\in\Lambda}\Big(\mathrm{Tr}_1[X_x(\gamma_{gs}+2\gamma_0)]\Big)^2\right)^{1/2}\left(\sum_{x\in\Lambda}\mathrm{Tr}_1[X_x(\mathbf{1}-\gamma_0)\gamma_{gs}(\mathbf{1}-\gamma_0)]\right)^{1/2} \\
&\geq -\mathrm{const}\sqrt{6}\left(\mathrm{Tr}_1\Big[\sum_{x\in\Lambda} X_x(\gamma_{gs}+2\gamma_0)\Big]\right)^{1/2}\left(\mathrm{Tr}_1\Big[\sum_{x\in\Lambda} X_x(\mathbf{1}-\gamma_0)\gamma_{gs}(\mathbf{1}-\gamma_0)\Big]\right)^{1/2} \quad (18)\\
&= -\mathrm{const}\, n^{1/2}\sqrt{\mathrm{Tr}_1[(\mathbf{1}-\gamma_0)\gamma_{gs}]} \ = \ -\mathrm{const}\, n^{1/2}\sqrt{\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_0]} \ . \quad (19)
\end{aligned}$$

In (18) we have used, that $\mathrm{Tr}_1[X_x] = 2$, which implies $\mathrm{Tr}_1[X_x(\gamma_{gs}+2\gamma_0)] \leq 3\mathrm{Tr}_1[X_x] \leq 6$, and in (19) we have used Equation 9, $(\mathbf{1}-\gamma_0) \leq 1$, and $\mathrm{Tr}_1[\gamma_0] = n|\Lambda| = \mathrm{Tr}_1[\gamma_{gs}]$. Similarly, we estimate $\sum_{x\in\Lambda} B_x$.

$$-\sum_{x\in\Lambda} B_x \ \leq \ \left(\sum_{x\in\Lambda}\big(\mathrm{Tr}_1[X_x\gamma_0]\big)^2\right)^{1/2}\left(\sum_{x\in\Lambda}\big(\mathrm{Tr}_1[X_x(\gamma_0-\gamma_{gs})]\big)^2\right)^{1/2} \ .$$

Now we use $\sum_{x\in\Lambda}\big(\mathrm{Tr}_1[X_x\gamma_0]\big)^2 \ \leq \ \sum_{x\in\Lambda}\mathrm{Tr}_1[X_x]\mathrm{Tr}_1[X_x\gamma_0] \ = \ 2n|\Lambda|$ and

$$\big(\mathrm{Tr}_1[X_x(\gamma_0-\gamma_{gs})]\big)^2 = \big(\mathrm{Tr}_1[X_x^{1/2}(\gamma_0-\gamma_{gs})X_x^{1/2}]\big)^2 \leq \mathrm{Tr}_1[X_x]\mathrm{Tr}_1[X_x(\gamma_0-\gamma_{gs})^2] \ .$$

Thus

$$\begin{aligned}
-\sum_{x\in\Lambda} B_x &\leq 2(n|\Lambda|)^{1/2}\left(\sum_{x\in\Lambda}\mathrm{Tr}_1[X_x(\gamma_0-\gamma_{gs})^2]\right)^{1/2} \\
&= 2(n|\Lambda|)^{1/2}\sqrt{\mathrm{Tr}_1[(\gamma_0-\gamma_{gs})^2]} \ = \ 2(n|\Lambda|)^{1/2}\sqrt{\mathrm{Tr}_1[\gamma_0+\gamma_{gs}] - 2\mathrm{Tr}_1[\gamma_0\gamma_{gs}]} \\
&= 2(2n|\Lambda|)^{1/2}\sqrt{\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_0]} \ . \quad (20)
\end{aligned}$$

Putting together (19) and (20) we arrive at the claim. ■

3 Bootstrapping

We have succeeded in estimating a two body problem in terms of a one body problem, namely, we have bounded $\triangle E(n)$ by

$$0 \geq \frac{\triangle E(n)}{|\Lambda|} \geq -\mathrm{const}\, U n^{1/2}\left(\frac{\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_0]}{|\Lambda|}\right)^{1/2} =: -\mathrm{const}\, U n^{1/2} A \ , \quad (21)$$

but still the 1-pdm γ_{gs} of the ground state is appearing in A. We will control A by a bootstrap argument. At this point we make use of the translational invariance of the hopping matrix $t_{xy} = t(x-y)$. For $U = 0$ we can diagonalize the Hamiltonian by Fourier transformation:

$$H = \sum_{k\in BZ} \mathrm{e}(k)\, c^\dagger_{k,\sigma} c_{k,\sigma} \quad ,$$

were $BZ = [-\pi,\pi)^d \cap \frac{\pi}{L}\mathbb{Z}^d$ is the Brillouin zone. Here $c^\dagger_{k,\sigma} = |\Lambda|^{-1/2}\sum_{x\in\Lambda}\exp(i\,k\cdot x)\,c^\dagger_{x,\sigma}$, and $\varphi_{k,\sigma} := c^\dagger_{k,\sigma}\Omega$ is the new ONB of $\mathcal{H}$. The single particle spectrum (dispersion relation) is given by

$$\mathrm{e}(k) := \hat{t}(k) = \sum_{r\in\Lambda} t(r)\, e^{i\,k\cdot r} \,. \tag{22}$$

If we fix the particle density per site $0 \le n \le 2$, such that $N = n|\Lambda|$ is even, we can write the ground state of the free Hubbard model as

$$\Psi_0 = \prod_{\mathrm{e}(k)\le\ \mathrm{e}_n\,,\,\sigma} c^\dagger_{k,\sigma}\Omega \,, \tag{23}$$

where e_n is the Fermi energy (defined by $\sum_{\,\mathrm{e}(k)\le\ \mathrm{e}_n} 1/|\Lambda| = n$). We shall henceforth assume N to be even without further mentioning. This does not cause a loss in generality for the theorem, because the error generated by passing from N to $N+1$ is

$$\frac{|E_{hf}|}{|\Lambda|} \;\le\; \mathrm{const}\,|\Lambda|^{-1}(1+U) \;\le\; \mathrm{const}\,\Big(U^{4/3} + U|\Lambda|^{-1/4}\Big) \,.$$

The **density of states** (DOS) is given by

$$\rho_d(\mathrm{e}_n) := \left.\frac{dN_d(E)}{dE}\right|_{E=\ \mathrm{e}_n} = \int\limits_{-\pi}^{\pi}\!\!\dots\!\!\int\limits_{-\pi}^{\pi} d^dk\;\; \delta(\mathrm{e}(k) - \mathrm{e}_n) \,, \tag{24}$$

where

$$N_d(E) = \int\limits_{-\pi}^{\pi}\!\!\dots\!\!\int\limits_{-\pi}^{\pi} d^dk\;\; \chi\left[\mathrm{e}(k)\le E\right] \,. \tag{25}$$

The DOS appears in the following bootstrap estimate:

Lemma 4 *There exist three constants $c_{1,2,3}$ such that for any $\varepsilon > 0$*

$$A^2 \;\le\; \frac{c_1 U n^{1/2} A}{\varepsilon} \;+\; c_2 \int\limits_{-\varepsilon-\frac{c_3}{L}}^{\frac{c_3}{L}} \rho_d(e_n + \lambda)\, d\lambda \,, \tag{26}$$

where A is defined in Equation 21 (recall that $2L$ is the width of the lattice).

Proof: Let

$$\begin{aligned}
\gamma_< \;&:=\; \sum_{\substack{\mathrm{e}(k)\le\ \mathrm{e}_n - \varepsilon \\ \sigma}} |\varphi_{k,\sigma}\rangle\langle\varphi_{k,\sigma}| \,, \text{ and} \\
\gamma_> \;&:=\; \sum_{\substack{\mathrm{e}_n - \varepsilon \le\ \mathrm{e}(k)\le\ \mathrm{e}_n \\ \sigma}} |\varphi_{k,\sigma}\rangle\langle\varphi_{k,\sigma}| \,.
\end{aligned}$$

We can write γ_0 in the Fourier transformed basis as

$$\gamma_0 \;=\; \sum_{\substack{\mathrm{e}(k)\le\ \mathrm{e}_n \\ \sigma}} |\varphi_{k,\sigma}\rangle\langle\varphi_{k,\sigma}| \;=\; \gamma_< + \gamma_> \,.$$

So we have that

$$\begin{aligned}
|\Lambda|A^2 &= \mathrm{Tr}[(\mathbf{1}-\gamma_{gs})\gamma_0] \\
&= \mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_<] + \mathrm{Tr}_1\Big[(\mathbf{1}-\gamma_{gs})^{1/2}\gamma_>(\mathbf{1}-\gamma_{gs})^{1/2}\Big] \\
&\leq \mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_<] + \mathrm{Tr}_1[\gamma_>] \; .
\end{aligned}$$

It holds that

$$\begin{aligned}
\gamma_< &\leq \sum_{\substack{\mathrm{e}(k)\leq\, \mathrm{e}_n-\varepsilon \\ \sigma}} \underbrace{\left(\frac{\mathrm{e}_n-\mathrm{e}(k)}{\varepsilon}\right)}_{\geq 1} |\varphi_{k,\sigma}\rangle\langle\varphi_{k,\sigma}| \\
&\leq \sum_{\substack{\mathrm{e}(k)\leq\, \mathrm{e}_n-\varepsilon \\ \sigma}} \left(\frac{\mathrm{e}_n-\mathrm{e}(k)}{\varepsilon}\right) |\varphi_{k,\sigma}\rangle\langle\varphi_{k,\sigma}| \;+ \sum_{\substack{\mathrm{e}_n-\varepsilon\leq\, \mathrm{e}(k)\leq\, \mathrm{e}_n \\ \sigma}} \underbrace{\left(\frac{\mathrm{e}_n-\mathrm{e}(k)}{\varepsilon}\right)}_{\geq 0} |\varphi_{k,\sigma}\rangle\langle\varphi_{k,\sigma}| \\
&= \left(\frac{\mathrm{e}_n-h}{\varepsilon}\right)\gamma_0 \; , \qquad (27)
\end{aligned}$$

(recall that h is the operator $t\otimes\mathbf{1}_{\mathbb{C}^2}$). Therefore

$$\begin{aligned}
\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_<] &\leq \frac{1}{\varepsilon}\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})(\mathrm{e}_n-h)\gamma_0] \\
&= \frac{1}{\varepsilon}\Big\{\mathrm{Tr}_1[(h-\mathrm{e}_n)(\gamma_{gs}-\gamma_0)] - \mathrm{Tr}_1[\gamma_{gs}(h-\mathrm{e}_n)(\mathbf{1}-\gamma_0)]\Big\} \\
&\leq \frac{1}{\varepsilon}\mathrm{Tr}_1[(h-\mathrm{e}_n)(\gamma_{gs}-\gamma_0)] \qquad (28) \\
&= \frac{1}{\varepsilon}\mathrm{Tr}_1[h(\gamma_{gs}-\gamma_0)] \; . \qquad (29)
\end{aligned}$$

To see why (28) holds it suffices to notice that $(h-\mathrm{e}_n)(\mathbf{1}-\gamma_0)\geq 0$. In (29) we have used that $\mathrm{Tr}_1[\gamma_{gs}-\gamma_0]=n|\Lambda|-n|\Lambda|=0$. So we get

$$\begin{aligned}
\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_<] &\leq \frac{1}{\varepsilon}\mathrm{Tr}_1[h(\gamma_{gs}-\gamma_0)] \\
&= \frac{1}{\varepsilon}\Big\{\mathrm{Tr}_1[h\gamma_{gs}] + \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}] - \underbrace{\mathcal{E}_{hf}(\gamma_0)}_{\geq E_{hf}(n)} \\
&\qquad + \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1}-\mathrm{Ex})(\gamma_0\otimes\gamma_0)] - \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}]\Big\} \\
&\leq \frac{1}{\varepsilon}\Big\{\triangle E(n) + \frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1}-\mathrm{Ex})(\gamma_0\otimes\gamma_0)] - \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}]\Big\} \qquad (30)
\end{aligned}$$

The last term in (30) we have already estimated with the help of the correlation inequalities. In the proof of Lemma 3 we had achieved that

$$\frac{1}{2}\mathrm{Tr}_2[V(\mathbf{1}-\mathrm{Ex})(\gamma_0\otimes\gamma_0)] - \frac{1}{2}\mathrm{Tr}_2[V\Gamma_{gs}] \leq \mathrm{const}\, U(n|\Lambda|)^{1/2}\big(\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_0]\big)^{1/2} \; .$$

Together with the fact that $\triangle E(n)\leq 0$ we arrive at

$$\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_<] \leq \frac{\mathrm{const}\, U\,(n|\Lambda|)^{1/2}\big(\mathrm{Tr}_1[(\mathbf{1}-\gamma_{gs})\gamma_0]\big)^{1/2}}{\varepsilon} \; . \qquad (31)$$

So we have proved one part of our claim. Note that until now every statement we made holds independently of the dimension of the lattice Λ. The dimension enters when we have to deal with the term $\mathrm{Tr}_1[\gamma_>]$. To estimate it, we compare it with its thermodynamic limit ($L \to \infty$). For each $k \in [-\pi, \pi]^d$, we may choose a unique $k_0(k) \in BZ$, such that $|(k - k_0)_j| \leq \pi/L$. Let

$$\tilde{\mathrm{e}}(k) := \mathrm{e}(k_0(k)) \ \text{ and } \ Q := [-\pi/2L, \pi/2L]^d \ ,$$

then it holds

$$|\mathrm{e}(k) - \tilde{\mathrm{e}}(k)| \leq \sqrt{(\pi/L)^2 d} \sup_{Q+k_0(k)} |\vec{\bigtriangledown}\mathrm{e}(k)| \ \leq \ \sqrt{(\pi/L)^2 d} \sup_{k \in BZ} |\vec{\bigtriangledown}\mathrm{e}(k)| \ =: \ \frac{c}{L} \ ,$$

and

$$\begin{aligned} \mathrm{Tr}_1[\gamma_>] \quad &= \quad \frac{|\Lambda|}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} \chi\left[\mathrm{e}_n - \varepsilon \leq \tilde{\mathrm{e}}(k) \leq \mathrm{e}_n\right] d^d k \\ &\leq \quad \frac{|\Lambda|}{(2\pi)^d} \int\limits_{[-\pi,\pi]^d} \chi\left[\mathrm{e}_n - \varepsilon - c/L \leq \mathrm{e}(k) \leq \mathrm{e}_n + c/L\right] d^d k \\ &= \quad \frac{|\Lambda|}{(2\pi)^d} \int\limits_{-\varepsilon - \frac{c}{L}}^{\frac{c}{L}} \rho_d(\lambda + \mathrm{e}_n) \, d\lambda \ . \qquad \blacksquare \end{aligned} \tag{32}$$

4 Proof of the Theorem

To prove the theorem we need an upper bound on the DOS so that we can make use of the bootstrap estimate (Lemma 4). We assume that the dispersion relation is a Morse function (i.e., a smooth function with finitely many nondegenerate critical points). This is a classical condition tracing back to Van Hove [7], and it implies the following facts:

- If $d \geq 3$ the DOS is finite.
- If $d = 2$ it holds
$$\rho_2(E) \ \leq \ \mathrm{const} \left|\ln|E_{crit} - E|\right| \ . \tag{33}$$
E_{crit} is the value of $\mathrm{e}(k)$ at the critical point with index 1 (the number of negative eigenvalues of the Hessian is 1). Without loss of generality we may consider the generic case: existence of exactly one critical energy for the two dimensional dispersion relation. The existence of more critical values does not alter our main result, except for possibly increasing the constants. Every smooth function on the torus has at least one critical point with index 1.

As an example for the two dimensional case consider nearest neighbour hopping for which the dispersion relation is $\mathrm{e}(k) = -2(\cos k_1 + \cos k_2)$. It has the four saddle points $(-\pi, 0)$, $(0, -\pi)$, $(\pi, 0)$, and $(0, \pi)$, lying on the Fermi-surface to the energy zero (corresponding to the density $n = 1$).

Proof of the Theorem: We deal with the simpler case $d \geq 3$ first. Let us collect the results of the last two chapters. We have that

$$\frac{\triangle E(n)}{|\Lambda|} \geq -\text{const } U n^{1/2} A \ , \tag{34}$$

(Lemma 3), where A is given by

$$A \ = \ \left(\frac{\mathrm{Tr}_1[(\mathbf{1} - \gamma_{gs})\gamma_0]}{|\Lambda|} \right)^{1/2} \ .$$

Furthermore (Lemma 4) there exist three constants $c_{1,2,3}$ such that for any $\varepsilon > 0$

$$\begin{aligned} A^2 \ &\leq \ \frac{c_1 U n^{1/2} A}{\varepsilon} \ + \ c_2 \int_{-\varepsilon - \frac{c_3}{L}}^{\frac{c_3}{L}} \rho_d(\mathrm{e}_n + \lambda) \, d\lambda \\ &\leq \ \frac{c_1 U n^{1/2} A}{\varepsilon} \ + \ c_4(\varepsilon + |\Lambda|^{-\frac{1}{d}}) \end{aligned} \tag{35}$$

(the constant c_4 is given by $c_4 = c_2 \sup\{\rho_d\} \max\{1, 4c_3\}$). We obtain

$$A \ \leq \ \frac{c_1 U n^{1/2}}{2\varepsilon} + \sqrt{\left(\frac{c_1 U n^{1/2}}{2\varepsilon} \right)^2 \ + \ c_4 \varepsilon \ \ + \ c_4 |\Lambda|^{-\frac{1}{2d}}} \tag{36}$$

from solving the quadratic equation corresponding to (35). We optimize in ε by choosing $\varepsilon = \text{const}\, n^{1/3} U^{2/3}$. Inserting this into (36) yields

$$A \ \leq \ \text{const } (n^{1/6} U^{1/3} + |\Lambda|^{-\frac{1}{2d}}) \ ,$$

and inserting this into (34) we get

$$\frac{\triangle E(n)}{|\Lambda|} \ \geq \ -\text{const } (n^{2/3} U^{4/3} + n^{1/2} U |\Lambda|^{-1/2d}) \ , \tag{37}$$

which is our claim for $d \geq 3$.

If the dimension is equal to two we have to consider the Van Hove singularity of the DOS at the critical energy. Let $\delta := E_{crit} - E$ and $x := \varepsilon + \frac{c_2}{L}$. We start again with the statement of Lemma 4

$$A^2 \ \leq \ \frac{c_1 U n^{1/2} A}{\varepsilon} \ + \ c_2 \underbrace{\int_{-\varepsilon - \frac{c_3}{L}}^{\varepsilon + \frac{c_3}{L}} \left| \ln |E_{crit} - E + \lambda| \right| \, d\lambda}_{=: F(\delta, x)} \ , \tag{38}$$

where we have inserted (33) and increased the length of the interval of integration by ε. We estimate $F(\delta, x)$ using Schwarz's inequality:

$$F(\delta, x) \ = \ \int_{-x}^{+x} \left| \ln |\delta + \lambda| \right| \, d\lambda \ \leq \ \sqrt{2x} \left(\int_{\delta - x}^{\delta + x} (\ln |\lambda|)^2 \, d\lambda \right)^{1/2} \ . \tag{39}$$

Without loss of generality we may assume that $\delta \geq 0$. If $x \geq \delta/2$

$$\int\limits_{\delta-x}^{\delta+x} (\ln|\lambda|)^2 \, d\lambda \;\leq\; 2\int\limits_{0}^{2x} (\ln\lambda)^2 \, d\lambda \;\leq\; \text{const}\, x(|\ln x|+1) \,. \tag{40}$$

If $x \leq \delta/2$ it is not possible to estimate the integral uniformly in δ, but

$$\begin{aligned} \int\limits_{\delta-x}^{\delta+x} (\ln|\lambda|)^2 \, d\lambda &\leq 2x \max\Big\{ \big(\ln(\delta/2)\big)^2, \big(\ln(3\delta/2)\big)^2 \Big\} \\ &\leq 2x\Big((|\ln x|+1)(\ln(3\delta/2+e)) \Big)^2 \,, \end{aligned} \tag{41}$$

where $\ln e = 1$. As δ is bounded by $\max\{|\mathrm{e}(k) - \mathrm{e}(k')|\}$ (the bandwidth), $\ln(3\delta/2+e)$ is bounded. Inserting (40) and (41) into (39) we arrive at

$$F(\delta, x) \;\leq\; \text{const}\, x(|\ln x|+1) \,. \tag{42}$$

It is easy to see that for all positive $\alpha_{1,2}$ there exists a constant c, such that

$$(\alpha_1+\alpha_2)|\ln(\alpha_1+\alpha_2)| \;\leq\; c\alpha_1|\ln(c\alpha_1)| + c\alpha_2|\ln(c\alpha_2)| \,.$$

This yields

$$F(\delta, x) \;\leq\; \text{const}\,\Big(\varepsilon(|\ln\varepsilon|+1) + |\Lambda|^{-1/2}(|\ln(|\Lambda|^{-1/2})|+1) \Big) \,, \tag{43}$$

(recall that $x = \varepsilon + c_2/L$). So we have estimated F in terms of ε and the lattice size. Inserting this into (38) we get

$$A^2 \;\leq\; \frac{c_1 U n^{1/2}}{\varepsilon} A + c_2\Big\{ \varepsilon(|\ln\varepsilon|+1) + |\Lambda|^{-1/2}(|\ln(|\Lambda|^{-1/2})|+1) \Big\} \,, \tag{44}$$

which implies that

$$A \;\leq\; \frac{c_1 U n^{1/2}}{2\varepsilon} + \sqrt{\left(\frac{c_1 U n^{1/2}}{2\varepsilon}\right)^2 + c_2\Big\{ \varepsilon(|\ln\varepsilon|+1) + |\Lambda|^{-1/2}(|\ln(|\Lambda|^{-1/2})|+1) \Big\}} \,. \tag{45}$$

Setting $\varepsilon = \text{const}\, n^{1/3} U^{2/3}$ like in the first part of our proof we arrive at

$$\begin{aligned} A &\leq \text{const}\,\Big(n^{1/6}U^{1/3} + n^{1/6}U^{1/3}\Big\{ \big(|\ln(n^{1/3}U^{2/3})|+1\big)^{1/2} + |\Lambda|^{-1/4}\big(|\ln(|\Lambda|^{-1/2})|+1\big)^{1/2} \Big\} \Big) \\ &\leq \text{const}\,\Big(n^{1/6}U^{1/3}(|\ln U|+1) + |\Lambda|^{-1/4}(|\ln(|\Lambda|^{-1/2})|+1) \Big) \,. \end{aligned}$$

So we finally arrive at the claim for $d = 2$ (feeding the last inequality into (34)):

$$\triangle E(n)/|\Lambda| \;\geq\; -\text{const}\,\Big(n^{2/3}U^{4/3}(|\ln U|+1) \;+\; n^{1/2}U|\Lambda|^{-1/4}(|\ln(|\Lambda|^{-1/2})|+1) \Big) \,. \quad \blacksquare$$

References

[1] V. Bach. Error bound for the Hartree-Fock energy of atoms and molecules. *Commun. Math. Phys.*__147__: 527-548, 1992.

[2] V. Bach. Applications for a Correlation Estimate for Fermions. *"Proceedings of the XIth International Congress of Mathematical Physics"; International Press*, 479-489, 1995.

[3] V. Bach, E.H. Lieb, J.P. Solovej. Generalized Hartree-Fock theory and the Hubbard model. *J.Stat.Phys.*__76__: 3-90, 1994.

[4] G.M. Graf, J.P. Solovej. A correlation estimate with applications to quantum systems with Coulomb interactions. *"The State of Matter - A volume dedicated to E.H. Lieb" World Scientific* (1994) *and Rev. Math. Phys.* **6**: 977-997, 1994.

[5] E.H. Lieb. Variational principle for many-fermion systems. *Phys.Rev.Lett.* **46**: 457-459, 1981.

[6] H.J. Schulz. Incommensurate Antiferromagnetism in the Two-Dimensional Hubbard Model. *Phys.Rev.Lett.*, **64**: 1445-1448, 1990.

[7] L. Van Hove. The Occurrence of Singularities in the Elastic Frequency Distribution of a Crystal. *Phys. Rev.* **89**: 1189-1193, 1953.