Bogolubov-Hartree-Fock Mean Field Theory for Neutron Stars and other Systems with Attractive Interactions

Volker Bach * Jürg Fröhlich † Lars Jonsson ‡

19-Dec-2008

Abstract

A simplification of Bogolubov-Hartree-Fock theory, which is a natural generalization of the traditional Hartree-Fock theory, is derived. This simplification allows to express the pairing interaction in terms of the one-particle density matrix for systems interacting by attractive pair potentials, like the Newtonian gravitational potential.

*Institut für Mathematik FB 08; Johannes Gutenberg-Universität; D-55099 Mainz; Germany; vbach@mathematik.uni-mainz.de

†Institut für Theoretische Physik; ETH Hönggerberg; CH-8044 Zürich; juerg@itp.phys.ethz.ch

‡School of Electrical Engineering; Royal Institute of Technology (KTH); Teknikringen 33; SE-100 44 Stockholm; ljonsson@kth.se

I Introduction

The purpose of this paper is the derivation of a simplification of Bogolubov-Hartree-Fock theory, which is a natural generalization of the traditional Hartree-Fock theory. This generalization is relevant for, and applies to, attractive pair interaction potentials, like the Newtonian gravitational potential.

In this section we start, however, at the origin of the Hartree-Fock approximation method in atomic and molecular physics. We recall Lieb's variational principle from which one obtains a natural link between Hartree-Fock theory and density matrices ρ on fermion Fock space via their one-particle density matrices γ_ρ. Among the preimages of γ_ρ under the map $\rho \mapsto \gamma_\rho$ is a distinguished density matrix which is quasifree and gauge-invariant, i.e., commutes with the particle number operator. The extension of the Hartree-Fock variational principle to *all* quasifree density matrices (not only the gauge-invariant ones) is what we call Bogolubov-Hartree-Fock theory. It yields a new map $\rho \mapsto (\gamma_\rho, \alpha_\rho)$ and an additional contribution to the energy depending on the *pairing* α_ρ. For repulsive pair interactions, this contribution is always positive and hence never visible after minimizing the energy functional, i.e., $\alpha = 0$ in this case. Thus, for repulsive systems, Bogolubov-Hartree-Fock theory reduces to Hartree-Fock theory. For attractive systems, however, pairing in the Bogolubov-Hartree-Fock functional possibly decreases the energy by an additional amount. All this has been described in [2].

The novelty of the present paper lies in the elimination of the pairing, as an additional degree of freedom, of the minmization problem. That is, for attractive systems of spin-$\frac{1}{2}$, we show that the pairing of the minimizer (γ, α) of the Bogolubov-Hartree-Fock functional assumes the form $\gamma = \hat{\gamma} \otimes \mathbf{1}$ and $\alpha = \sqrt{\hat{\gamma} - \hat{\gamma}^2} \otimes \sigma_x$, where $\hat{\gamma}$ is a real one-particle density matrix of half the particle number, and σ_x is the first Pauli matrix. We consider our work as the first part of a continuation of [9, 8, 7].

In spite of what this description might suggest, our result does not require the existence of a minimizer. This is important because the lack of weak lower semicontinuity makes its proof a difficult task. Remarkably, for the model of main interest - pseudorelativistic fermions interacting by Newtonain gravitational forces - the existence of minimizers and their characterization has been achieved in [11]. Indeed, [11] provides a series of results complementing our present paper.

Bogolubov-Hartree-Fock theory is not the only possible generalization of Hartree-Fock theory. The Müller functional, introduced and analyzed in [6], constitutes an interesting alternative generalization of Hartree-Fock theory, which has also been inspiring for the present paper.

I.1 The Hartree-Fock Approximation

One of the most important approximation methods in quantum mechanics of many particles is the Hartree-Fock (HF) approximation. It defines a qualitatively accurate theory for molecular binding, and much of our intuitive understanding of atomic and molecular physics is based on it. Its ground state, the HF ground state, is defined by restricting the Raleigh-Ritz variational principle for the lowest eigenvalue of the Hamiltonian to a minimization over Slater determinants

$$\Phi \;=\; \varphi_1 \wedge \cdots \wedge \varphi_N, \qquad \langle \varphi_i | \varphi_j \rangle = \delta_{ij}, \tag{I.1}$$

i.e., antisymmetric products of N orthonormal one-fermion wave functions (which we henceforth refer to as *orbitals*). The orbitals $\{\varphi_i\}_{i=1}^N \subseteq \mathfrak{h}$ are elements of the (separable) one-particle Hilbert space $\mathfrak{h}$, which we assume to be an L^2-space, $\mathfrak{h} = L^2(M, dx)$ of some set M, without loss of generality. More specifically, if the Hamiltonian of the system of N fermions is given as

$$H_N \;:=\; \sum_{i=1}^N h_i \;+\; \sum_{1 \le i < j \le N} W_{i,j}, \tag{I.2}$$

where h is a one-particle operator (acting on $\mathfrak{h}$), and $W_{i,j} = W(x_i, x_j) \ge 0$ is a repulsive pair interaction potential (acting on $\mathfrak{h} \otimes \mathfrak{h}$), and if Φ is a Slater determinant as in (I.1) then

$$\langle \Phi \mid H_N \Phi \rangle \;=\; \sum_{i=1}^N \langle \varphi_i | h \varphi_i \rangle \;+\; \frac{1}{2} \int \Big\{ \rho_\Phi(x) \rho_\Phi(y) - |\gamma_\Phi(x,y)|^2 \Big\} W(x,y)\, dx\, dy, \tag{I.3}$$

where $\gamma_\Phi(x,y) := \sum_{i=1}^N \varphi_i(x) \overline{\varphi_i(y)}$ is the one-particle density matrix (1-pdm) corresponding to Φ and $\rho_\Phi(x) := \sum_{i=1}^N |\varphi_i(x)|^2$ its one-particle density. The HF energy is defined as

$$E_{HF}(N) \;:=\; \inf \big\{ \langle \Phi \mid H_N \Phi \rangle \;\big|\; \Phi \text{ as in (I.1)} \big\}. \tag{I.4}$$

The HF energy is always an upper bound to the quantum mechanical ground state energy. To appreciate this fact, we note that only few approximation methods in quantum mechanics have such a clear and simple mathematical relation to fundamental quantum mechanics.

I.2 Lieb's Variational Principle

From a mathematical point of view, however, the Slater determinants of N orbitals define a rather unstructured subset of the Hilbert space of all N-fermion wave functions: In general, a sum of two Slater determinants is not a Slater determinant, and a convex combination of projections onto Slater determinants is not a projection onto a Slater determinant, either.

Analogously, the 1-pdm $\gamma_\Phi = \sum_{i=1}^N |\varphi_i\rangle\langle\varphi_i|$ of an N-orbital Slater determinant $\Phi = \varphi_1 \wedge \cdots \wedge \varphi_N$ is a rank-N orthogonal projection, and convex combinations of two such projections never yield a projection unless they coincide. Yet, 1-pdm provide a convenient formulation of the HF variational principle, which we describe next.

We introduce the Hartree-Fock energy functional

$$\mathcal{E}_{HF}(\gamma) \;:=\; \mathrm{Tr}\{h\,\gamma\} \;+\; \frac{1}{2}\int \Big\{\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x,y)|^2\Big\} W(x,y)\,dx\,dy, \quad \text{(I.5)}$$

for any 1-pdm $0 \le \gamma \le 1$, i.e., any self-adjoint trace-class operator $\gamma = \sum_{i=1}^\infty \lambda_i\,|\psi_i\rangle\langle\psi_i|$ with orthonormal eigenvectors ψ_i and corresponding eigenvalues λ_i between 0 and 1, where $\gamma(x,y) = \sum_{i=1}^\infty \lambda_i\,\psi_i(x)\overline{\psi_i(y)}$ and $\rho_\gamma(x) := \gamma(x,x)$. Since

$$\mathcal{E}_{HF}(\gamma_\Phi) \;=\; \langle\Phi \mid H_N\Phi\rangle, \quad \text{(I.6)}$$

if γ_Φ corresponds to a Slater determinant Φ, the definition (I.4) of the HF energy can be rewritten in terms of 1-pdm as

$$E_{HF}(N) \;=\; \inf\big\{\mathcal{E}_{HF}(\gamma) \;\big|\; \gamma = \gamma^* = \gamma^2, \;\; \mathrm{Tr}(\gamma) = N\big\}. \quad \text{(I.7)}$$

An important observation of Lieb [13] (for a simple proof see also [1]) is that for repulsive interaction potentials, $W(x,y) \ge 0$, the condition in (I.7) on γ to be a projection can be relaxed to $0 \le \gamma \le 1$ without changing the infimum,

$$E_{HF}(N) \;=\; \inf\big\{\mathcal{E}_{HF}(\gamma) \;\big|\; 0 \le \gamma \le 1, \;\; \mathrm{Tr}(\gamma) = N\big\}. \quad \text{(I.8)}$$

Hence, the HF energy is the minimal expectation of the HF energy functional varied over all 1-pdm of trace N. Lieb's observation has two important consequences. First, the set of 1-pdm is convex, i.e., closed under taking convex combinations, which is an important advantage in the calculus of variations. Secondly, 1-pdm can be naturally defined not only for Slater determinants, but for any many-particle wave function and even for any mixed state with particle number expectation value N.

We remark that much of the conclusion of Lieb's variational principle (I.8) holds true for general potentials (repulsive, attractive, or even indefinite) under the additional assumption of existence of a minimizer. More precisely, if the infimum on the right side of (I.8) is actually a minimum then the effective Hamiltonian arising from the linearization of the HF functional possesses at least N eigenvalues $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_N$ below its essential spectrum, and the minimizing 1-pdm is the orthogonal projection onto the spectral subspace correpsonding to $\lambda_1, \ldots, \lambda_{N-1}$ plus some operator of trace one whose range is contained in the spectral subspace of λ_N. We prove this assertion in Lemma 3 below.

I.3 Fock Space and Second Quantization

The fact that any pure or mixed state yields a 1-pdm leads us to the question whether the converse is true, as well, in the sense that, for a given 1-pdm of trace N, there is a naturally defined state with particle number expectation value N. The answer is affirmative as (generalized) 1-pdm can be naturally identified with quasifree states. To describe this identification, we consider the fermion Fock space $\mathfrak{F}[\mathfrak{h}]$ over $\mathfrak{h}$,

$$\mathfrak{F} \;\equiv\; \mathfrak{F}[\mathfrak{h}] \;=\; \oplus_{N=0}^{\infty} \mathfrak{H}^{(N)}, \tag{I.9}$$

where $\mathfrak{H}^{(N)} = \bigwedge_{n=1}^{N} \mathfrak{h}$ is the N-fermion Hilbert space, $\mathfrak{H}^{(0)} = \mathbb{C}\,\Omega$ is the vacuum sector, and Ω the (normalized) vacuum vector. It is convenient to represent $\mathfrak{F}$ in terms of creation and annihilation operators. For given orbitals $\varphi_1, \ldots, \varphi_N \in \mathfrak{h}$, we identify

$$\varphi_1 \wedge \cdots \wedge \varphi_N \;=:\; c^\dagger(\varphi_1)\, c^\dagger(\varphi_2) \cdots c^\dagger(\varphi_N)\, \Omega \;=:\; c^\dagger(\varphi_1)\, \big[c^\dagger(\varphi_2) \cdots c^\dagger(\varphi_N)\, \Omega\big], \tag{I.10}$$

which recursively defines the action of the creation operators $c^\dagger(\varphi)$ on vectors in $\mathfrak{F}$. Hence,

$$\mathfrak{H}^{(N)} \;=\; \overline{\mathrm{span}\,\big\{c^\dagger(\varphi_1) \cdots c^\dagger(\varphi_N)\, \Omega \;\big|\; \varphi_1, \ldots, \varphi_N \in \mathfrak{h}\big\}}, \tag{I.11}$$

and $\mathfrak{F}$ is the norm closure of polynomials in creation operators acting on the vacuum vector. Together with their adjoints $c(\varphi) = \big(c^\dagger(\varphi)\big)^*$, the annihilation operators, the creation operators constitute the Fock representation of the canonical anticommutation relations (CAR)

$$\begin{aligned}
\forall\, \varphi, \psi \in \mathfrak{h}: \quad \{c(\varphi), c(\psi)\} \;&=\; \{c^\dagger(\varphi), c^\dagger(\psi)\} \;=\; 0, \\
\{c(\varphi), c^\dagger(\psi)\} \;&=\; \langle \varphi | \psi \rangle, \\
c(\varphi)\Omega \;&=\; 0,
\end{aligned} \tag{I.12}$$

where the anticommutator is defined as $\{A, B\} := AB + BA$. Note that the CAR imply, in particular, that fermion creation and annihilation operators are bounded, namely $\|c^\dagger(\varphi)\|, \|c(\varphi)\| \leq \|\varphi\|$. In terms of creation and annihilation operators, the number operator $\mathbb{N}$ and the second-quantized

Hamiltonian $\mathbb{H}$ can be written as

$$\mathbb{N} \;:=\; \bigoplus_{N=0}^{\infty} (N \cdot \mathbf{1}_{\mathfrak{H}^{(N)}}) \;=\; \sum_{k=1}^{\infty} c_k^\dagger \, c_k, \tag{I.13}$$

$$\begin{aligned} \mathbb{H} \;&:=\; \bigoplus_{N=0}^{\infty} H_N \\ &=\; \sum_{k,\ell=1}^{\infty} h_{k,\ell}\, c_k^\dagger \, c_\ell \;+\; \sum_{k,\ell,m,n=1}^{\infty} W_{k,\ell;m,n}\, c_\ell^\dagger \, c_k^\dagger \, c_m \, c_n, \end{aligned} \tag{I.14}$$

where $\{\varphi_k\}_{k=1}^{\infty} \subseteq \mathfrak{h}$ is an orthonormal basis, $h_{k,\ell} := \langle \varphi_k | h \varphi_\ell \rangle$, $W_{k,\ell;m,n} := \langle \varphi_k \otimes \varphi_\ell | W \varphi_m \otimes \varphi_n \rangle$, $c_k^\dagger := c^\dagger(\varphi_k)$, $c_\ell := c(\varphi_\ell)$, for all $k, \ell, m, n \in \mathbb{N}$. Moreover, we assume that $\{\varphi_k\}_{k=1}^{\infty}$ is contained in the domain $\mathrm{dom}(h)$ of the self-adjoint one-particle operator h and that W is a relatively bounded perturbation of $h \otimes \mathbf{1} + \mathbf{1} \otimes h$. (A typical example for such an orthonormal basis $\{\varphi_k\}_{k=1}^{\infty} \subseteq \mathrm{dom}(h)$ in case that $\mathbb{H}$ is the second-quantized Hamiltonian of an atom or a molecule is the family of Hermite functions, the eigenfunctions of the harmonic oscillator.) Note that, though $\mathbb{H}$ is defined on finite vectors (i.e., polynomials in $c_k^\dagger$ applied to Ω), we do not claim $\mathbb{H}$ to be essentially selfadjoint on this domain.

I.4 States and Generalized One-Particle Density Matrices

A state $\omega \in \mathfrak{M}^*$ is a positive, linear, and normalized functional on a C^*-algebra $\mathfrak{M} \subseteq \mathcal{B}[\mathfrak{F}]$. The set

$$\mathfrak{S} \;:=\; \big\{ \omega \in \mathfrak{M}^* \;\big|\; \forall\, A \in \mathfrak{M},\; A \geq 0 :\; \omega(A) \geq 0, \;\; \omega(\mathbf{1}) = 1 \big\} \tag{I.15}$$

of all states is a convex subset of $\mathfrak{M}^*$. In this context, we choose $\mathfrak{M}$ to be the C^*-algebra generated by all polynomials in $c^\dagger(\varphi)$ and $c(\varphi)$ of even degree, where $\varphi \in \mathfrak{h}$. Then we may define the generalized one-particle density matrix (g1-pdm) $\Gamma_\omega \in \mathcal{B}[\mathfrak{h} \oplus \mathfrak{h}]$ for a given state $\omega \in \mathfrak{S}$ by

$$\left\langle \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \,\middle|\, \Gamma_\omega \, \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \right\rangle \;:=\; \omega\Big([c^\dagger(\psi_1) + c(\bar{\psi}_2)]\, [c^\dagger(\bar{\varphi}_1) + c(\varphi_2)] \Big), \tag{I.16}$$

where $\bar{\varphi}$ and $\bar{\psi}$ denote the complex conjugates of $\varphi : M \to \mathbb{C}$ and $\psi : M \to \mathbb{C}$, respectively, ensuring that (I.16) is linear in $\psi_1 \oplus \psi_2$ and antilinear in $\varphi_1 \oplus \varphi_2$. Note that complex conjugation is basis dependent, i.e., makes use of our assumption that $\mathfrak{h} = L^2(M)$. We could avoid this assumption and work with an abstract separable complex Hilbert space $\mathfrak{h}$ equipped with an antilinear involution J, but we choose not to do so. Due to the CAR, we obviously have

$$0 \;\leq\; \Gamma_\omega \;\leq\; \mathbf{1}, \tag{I.17}$$

which is equivalent to

$$\Gamma_\omega - \Gamma_\omega^2 \;\geq\; 0, \tag{I.18}$$

by functional calculus. It is convenient to represent the g1-pdm as a 2×2-matrix whose entries involve two bounded operators $\gamma_\omega, \alpha_\omega \in \mathcal{B}[\mathfrak{h}]$ defined by

$$\langle \psi | \gamma_\omega \, \varphi \rangle \;:=\; \omega\big(c^\dagger(\varphi)\, c(\psi) \big), \tag{I.19}$$

$$\langle \psi | \alpha_\omega \, \varphi \rangle \;:=\; \omega\big(c(\bar{\varphi})\, c(\psi) \big). \tag{I.20}$$

Namely, $\Gamma_\omega \;=\; \Gamma[\gamma_\omega, \alpha_\omega]$, with

$$\Gamma[\gamma, \alpha] \;=\; \begin{pmatrix} \gamma & \alpha \\ \alpha^* & \mathbf{1} - \overline{\gamma} \end{pmatrix}, \tag{I.21}$$

where, for a given $A \in \mathcal{B}[\mathfrak{h}]$, we define the complex conjugate $\overline{A} \in \mathcal{B}[\mathfrak{h}]$ in the obvious way: $\langle \psi | \overline{A}\, \varphi \rangle := \overline{\langle \psi | A\, \varphi \rangle}$. We observe that γ_ω is self-adjoint and α_ω is antisymmetric,

$$\gamma_\omega = \gamma_\omega^*, \qquad \alpha_\omega = -\alpha_\omega^T, \tag{I.22}$$

where α_ω^T is the transpose of α_ω, i.e., $\langle \psi | \alpha_\omega^T \, \varphi \rangle := \langle \bar{\varphi} | \alpha_\omega \, \bar{\psi} \rangle$. Note that (I.18) is equivalent to

$$\begin{pmatrix} \gamma_\omega - \gamma_\omega^2 - \alpha_\omega \alpha_\omega^* & (\mathbf{1} - \gamma_\omega)\alpha_\omega + \alpha_\omega(\mathbf{1} - \bar{\gamma}_\omega) \\ \alpha_\omega^*(\mathbf{1} - \bar{\gamma}_\omega) + (\mathbf{1} - \gamma_\omega)\alpha_\omega^* & \bar{\gamma}_\omega - \bar{\gamma}_\omega^2 - \alpha_\omega^* \alpha_\omega \end{pmatrix} \;=\; \Gamma_\omega - \Gamma_\omega^2 \;\geq\; 0, \tag{I.23}$$

which implies that

$$\gamma_\omega \;\geq\; \gamma_\omega^2 + \alpha_\omega \alpha_\omega^*. \tag{I.24}$$

Further note that if $\omega = |\Phi\rangle\langle\Phi|$ is the state corresponding to a Slater determinant Φ then γ_ω is its 1-pdm.

We collect the g1-pdm in the (convex) set

$$\mathfrak{A} \;:=\; \big\{ \Gamma = \Gamma(\gamma, \alpha) \in \mathcal{B}(\mathfrak{h} \oplus \mathfrak{h}) \;\big|\; 0 \leq \Gamma \leq \mathbf{1},\; \gamma = \gamma^*,\; \alpha = -\alpha^T \big\}. \tag{I.25}$$

Note that the 1-pdm of Sect. I.2 are special cases of g1-pdm, namely

$$\mathfrak{A}^{(0)} \;:=\; \big\{ \Gamma(\gamma, 0) \in \mathcal{B}(\mathfrak{h} \oplus \mathfrak{h}) \;\big|\; 0 \leq \gamma = \gamma^* \leq \mathbf{1} \big\} \;\subseteq\; \mathfrak{A}, \tag{I.26}$$

which we call the gauge invariant 1-pdm. Moreover, if $\Gamma(\gamma, \alpha) \in \mathfrak{A}$ then $0 \leq \gamma \leq \mathbf{1}$ and hence $\Gamma(\gamma, 0) \in \mathfrak{A}^{(0)}$. Further noticing that

$$\omega(\mathbb{N}) \;=\; \mathrm{Tr}(\gamma_\omega), \tag{I.27}$$

we define, for any $N < \infty$,

$$\mathfrak{A}_N \;:=\; \big\{\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A} \;\big|\; \mathrm{Tr}(\gamma) = N\big\}, \tag{I.28}$$

$$\mathfrak{A}_N^{(0)} \;:=\; \big\{\Gamma = \Gamma(\gamma, 0) \in \mathfrak{A}^{(0)} \;\big|\; \mathrm{Tr}(\gamma) = N\big\}, \tag{I.29}$$

the g1-pdm and gauge invariant 1-pdm, respectively, of particle number N.

Finally, we embed the g1-pdm in a suitable Banach space. To this end, we introduce

$$\mathcal{Y} \;:=\; \big\{G \in \mathcal{B}(\mathfrak{h} \oplus \mathfrak{h}) \;\big|\; G = G^*, \quad \overline{\tau G \tau^*} = G\big\}, \tag{I.30}$$

topologized by the operator norm $\| \cdot \|_{\mathrm{op}}$ on $\mathfrak{h} \oplus \mathfrak{h}$ and where

$$\tau \;:=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{I.31}$$

Obviously, $\mathcal{Y} \subset \mathcal{B}(\mathfrak{h} \oplus \mathfrak{h})$ is a closed subspace, which contains $\mathfrak{A}$ and $\mathfrak{A}_N$ as convex subsets in a natural way. Namely,

$$\mathfrak{A} \;=\; \Big\{\Gamma \in \mathcal{B}(\mathfrak{h} \oplus \mathfrak{h}) \;\Big|\; \big(\Gamma - \tfrac{1}{2}\big) \in \mathcal{Y}, \quad \big\|\Gamma - \tfrac{1}{2}\big\|_{\mathrm{op}} \leq \tfrac{1}{2}, \quad p^*\Gamma p \in \mathcal{L}^1(\mathfrak{h})\Big\}, \tag{I.32}$$

$$\mathfrak{A}_N \;=\; \Big\{\Gamma \in \mathcal{B}(\mathfrak{h} \oplus \mathfrak{h}) \;\Big|\; \big(\Gamma - \tfrac{1}{2}\big) \in \mathcal{Y}, \quad \big\|\Gamma - \tfrac{1}{2}\big\|_{\mathrm{op}} \leq \tfrac{1}{2}, \quad \mathrm{Tr}[p^*\Gamma p] = N\Big\}, \tag{I.33}$$

where $p : \mathfrak{h} \oplus \mathfrak{h} \to \mathfrak{h}$, $\psi \oplus \varphi \mapsto \psi$.

I.5 Density Matrices and Quasifree States

We introduce two important subclasses of states (a detailed account on the theory of operator algebras can be found in [3, 4]). The first is the set of states on $\mathcal{B}[\mathfrak{F}]$ (rather than a C^*-subalgebra of it). This is the set $\mathscr{L}^1_+[\mathfrak{F}] \subseteq \mathcal{B}[\mathfrak{F}]$ of density matrices, i.e., positive operators $\rho = \rho^* \geq 0$ of trace one, $\mathrm{Tr}\{\rho\} = 1$. Namely, given a density matrix ρ, the map $A \mapsto \omega_\rho(A) := \mathrm{Tr}\{\rho A\}$, with $A \in \mathcal{B}[\mathfrak{F}]$, defines a state. This map yields all bounded linear functionals on $\mathcal{B}[\mathfrak{F}]$, and the trace class operators are the dual of $\mathcal{B}[\mathfrak{F}]$. In particular, the projection $|\Psi\rangle\langle\Psi|$ onto a normalized vector $\Psi \in \mathfrak{F}$ is a density matrix (of rank one), and the corresponding state is called pure.

If $\mathfrak{M} \subset \mathcal{B}[\mathfrak{F}]$ is a subalgebra, then $\mathfrak{M}^* \supset \mathcal{B}[\mathfrak{F}]^*$, by duality. In general, the set of states given by density matrices is rather small, but it contains all states with finite particle number expectation. More precisely, if $\omega \in \mathfrak{S}_N$, where

$$\mathfrak{S}_N \ := \ \big\{ \omega \in \mathfrak{S} \ \big| \ \ \mathrm{Tr}(\gamma_\omega) = N \ \big\}, \tag{I.34}$$

denotes the (convex) set of states of particle number expectation $N < \infty$, then there exists a density matrix $\rho \in \mathscr{L}^1_+[\mathfrak{F}]$ such that $\omega = \omega_\rho$. In this paper we always assume the finiteness of the particle number expectation, so there is no need to distinguish between states and density matrices. Therefore, we may henceforth write Γ_ρ, γ_ρ, and α_ρ instead of Γ_{ω_ρ}, γ_{ω_ρ}, and α_{ω_ρ}, respectively, provided the context is clear.

The second important subclass of states is the family of quasifree states. A state $\omega \in \mathfrak{S}$ is called quasifree iff all truncated correlation functions except the 2-point function vanish. That is, given $2n \in 2\mathbb{N}$ orbitals $\varphi_1, \ldots \varphi_{2n}$ and $2n$ creation and annihilation operators $C_1, \ldots, C_{2n} \in \{c^\dagger(\varphi_1), c(\varphi_1), \ldots, c^\dagger(\varphi_{2n}), c(\varphi_{2n})\}$, we have $\omega(C_1 \cdots C_{2n-1}) = 0$ and

$$\omega(C_1 \cdots C_{2n}) \ = \ \sum_{\pi \in \mathcal{P}_{2n}} (-1)^\pi \, \omega(C_{\pi(1)} C_{\pi(2)}) \cdots \omega(C_{\pi(2n-1)} C_{\pi(2n)}), \tag{I.35}$$

where $\mathcal{P}_{2n} \subseteq \mathcal{S}_{2n}$ denotes the set of pairings of $2n$ elements, i.e., all permutations $\pi \in \mathcal{S}_{2n}$, for which $\pi(1) < \pi(3) < \ldots < \pi(2n-1)$ and $\pi(2) < \pi(4) < \ldots < \pi(2n)$. In particular,

$$\omega(C_1 C_2 C_3 C_4) \ = \ \omega(C_1 C_4)\omega(C_2 C_3) \, - \, \omega(C_1 C_3)\omega(C_2 C_4) \, + \, \omega(C_1 C_2)\omega(C_3 C_4). \tag{I.36}$$

We collect all quasifree states in the set $\mathfrak{Q} \subseteq \mathfrak{S}$, and we set $\mathfrak{Q}_N := \mathfrak{Q} \cap \mathfrak{S}_N$.

We list a few facts about quasifree states and density matrices.

- The state ω_ρ corresponding to a density matrix $\rho \in \mathscr{L}^1[\mathfrak{F}]$ is quasifree iff there exists a nonnegative selfadjoint operator $Q \geq 0$ on a subspace $\Pi\mathcal{F} \subseteq \mathcal{F}$ which is quadratic in the creation and annihilation operators and such that $\rho = \Pi \, \exp[-Q] \, \Pi$, where Π is the orthogonal projection onto $\Pi\mathcal{F}$. In fact, Π itself is the second quantization of a projection, and hence there is a quadratic selfadjoint operator $\widetilde{Q} \geq 0$ on $\mathcal{F}$, which commutes with Q and is such that $\Pi = \lim_{\beta\to\infty} \exp[-\beta\widetilde{Q}]$. In particular, we have

$$\begin{aligned} \mathfrak{Q}_N \ &= \ \big\{ \omega_\rho \ \big| \ \ \rho = \Pi \exp[-Q]\Pi, \ \ Q = Q^* = \Pi Q \Pi \text{ is quadratic}, \\ &\qquad \Pi = \Pi^* = \Pi^2, \ \ \mathrm{Tr}(\Pi \, e^{-Q}) = 1, \ \ \mathrm{Tr}(\mathbb{N} \, \Pi \, e^{-Q}) = N \big\}. \end{aligned} \tag{I.37}$$

- For every admissible 1-pdm $\Gamma = \Gamma[\gamma, \alpha] \in \mathfrak{A}$, there is a quasifree state $\omega \in \mathcal{Q}$ such that $\Gamma = \Gamma_\omega$, and in fact the maps

$$\mathfrak{Q} \rightarrow \mathfrak{A} \quad , \quad \omega \mapsto \Gamma_\omega, \tag{I.38}$$

$$\mathfrak{Q}_N \rightarrow \mathfrak{A}_N \quad , \quad \omega \mapsto \Gamma_\omega \tag{I.39}$$

  are bijections.

- There is also an obvious equivalence between $\alpha_\omega = 0$ and the corresponding density matrix to be gauge invariant,

$$\Big\{\rho \in \mathfrak{S}_N, \quad [\mathbb{N}, \rho] = 0\Big\} \quad \Leftrightarrow \quad \Big\{\Gamma_\rho \in \mathfrak{A}_N^{(0)}\Big\}. \tag{I.40}$$

II Definition and General Properties of the BHF Approximation

In this section we specify conditions on the one-body operator h, the pair interaction W, and the g1-pdm $\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A}$ that ensure the existence of the Bogolubov-Hartree-Fock (BHF) functional evaluated at Γ, its semi-boundedness, and other general properties.

Dealing with possibly unbounded operators h and W, we need to specify their domains of definition. Given the one-particle Hilbert space $\mathfrak{h} = L^2(M)$ of square-integrable functions on a configuration space M, we assume that the one-body Hamiltonian $h : \mathfrak{d} \times \mathfrak{d} \to \mathbb{C}$ is defined as a closed quadratic form on a dense form domain $\mathfrak{d} \subseteq \mathfrak{h}$, bounded below by

$$h \;\geq\; e_0 \cdot \mathbf{1}, \qquad \text{for some } e_0 \;>\; -\infty. \tag{II.41}$$

We assume W to be either positive or negative. That is, $W = V$ or $W = -V$, where $V : (\mathfrak{d} \otimes \mathfrak{d}) \times (\mathfrak{d} \otimes \mathfrak{d}) \to \mathbb{C}$ is a nonnegative quadratic form, with domain containing $(\mathfrak{d} \otimes \mathfrak{d}) \subseteq (\mathfrak{h} \otimes \mathfrak{h})$, which is a relatively form-bounded perturbation of $h \otimes \mathbf{1} + \mathbf{1} \otimes h$, with relative bound $a < \infty$. That is, there exist $a < \infty$ and $b \in [1 + |e_0|, \infty)$ such that, for all $\psi \in \mathfrak{d} \otimes \mathfrak{d}$,

$$0 \;\leq\; \langle \psi |\, V \psi \rangle \;\leq\; a \,\langle \psi |\, \{(h+b) \otimes \mathbf{1} + \mathbf{1} \otimes (h+b)\}\, \psi \rangle. \tag{II.42}$$

As a consequence, the form domain of $H_N = \sum_{i=1}^N h_i \pm \sum_{1 \leq i < j \leq N} V_{i,j}$ [recall (I.2)] contains $\bigwedge_{n=1}^N \mathfrak{d} \subseteq \mathfrak{H}^{(N)}$, for any finite particle number $N < \infty$.

We transcribe domain requirements to 1-pdm and g1-pdm, by assuming that $\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A}_N(h)$, where

$$\mathfrak{A}_N(h) \;:=\; \big\{ \Gamma(\gamma, \alpha) \in \mathfrak{A}_N \;\big|\; \|\gamma\|_{h,V} := \mathrm{Tr}\{(h+b)\gamma\} < \infty \big\} \tag{II.43}$$

is the set of g1-pdm with finite one-body expectation and of particle number N. Note that $\mathfrak{A}_N(h)$ is a convex subset of the Banach space $\mathcal{Y}$ introduced in (I.30) and that Condition (II.43) is a statement about the expectation value of h and $\mathbf{1}$ w.r.t. γ. A state whose g1-pdm is Γ may well have nonvanishing components in $\mathfrak{H}^{(K)}$ even for $K \gg N$.

To define the BHF functional it is convenient to introduce

$$D(f,g) \;:=\; \int f(x)\, g(y)\, V(x,y)\, dx\, dy, \tag{II.44}$$

$$P(\eta) \;:=\; \int |\eta(x,y)|^2\, V(x,y)\, dx\, dy, \tag{II.45}$$

for all $f, g \in L^1(M; \mathbb{R}_0^+)$ and all $\eta \in L^2(M \times M)$, for which these integrals are finite.

Lemma 1 *Suppose that $V \geq 0$ is relative h-form bounded with relative bound $a < \infty$, let $N \in \mathbb{N}$ and $\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A}_N(h)$. Then the following assertions hold true*

(i)

$$
\begin{aligned}
D(\rho_\gamma, \rho_\gamma) &\leq 2a\, N \,\mathrm{Tr}\{(h+b)\gamma\}, && \text{(II.46)}\\
P(\gamma) + P(\alpha) &\leq 2a \,\mathrm{Tr}\{(h+b)\gamma\}, && \text{(II.47)}
\end{aligned}
$$

In particular, the repulsive (+) and attractive (-) BHF-functional, resp.,

$$
\begin{aligned}
\mathcal{E}^{\pm}_{BHF}(\Gamma) &:= \mathrm{Tr}\{h\,\gamma\} \pm \frac{1}{2}\Big\{ D(\rho_\gamma, \rho_\gamma) - P(\gamma) + P(\alpha) \Big\} \qquad \text{(II.48)}\\
&= \mathrm{Tr}\{h\,\gamma\} \pm \frac{1}{2} \int \Big\{ \rho_\gamma(x)\rho_\gamma(y) - |\gamma(x,y)|^2 + |\alpha(x,y)|^2 \Big\} V(x,y)\, dx\, dy
\end{aligned}
$$

exists, $\mathcal{E}^{\pm}_{BHF} : \mathfrak{A}_N(h) \to \mathbb{R}$.

(ii)

$$
P(\gamma) \leq D(\rho_\gamma, \rho_\gamma), \qquad \text{(II.49)}
$$

and the repulsive BHF-functional $\mathcal{E}^{+}_{BHF}$ is bounded below and coercive on $\mathfrak{A}_N(h)$.

(iii) If $a < (N+1)^{-1}$ then the attractive BHF-functional $\mathcal{E}^{-}_{BHF}$ is bounded below and coercive on $\mathfrak{A}_N(h)$.

Before we give the proof of Lemma 1, we point out that (iii) yields only a poor estimate on a. By using Lieb-Thirring or Daubechies inequalities as in [9], it is possible to establish the semiboundeness and coercivity of $\mathcal{E}^{-}_{BHF}$ on $\mathfrak{A}_N(h)$ for pseudorelativistic fermions with attractive graviational interaction potentials, provided $a < cN^{-2/3}$, which is the natural semiclassical bound in this context, see [12].

Proof:

(i) If $\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A}_N(h)$ then α is Hilbert-Schmidt, and we may identify the integral kernel $\alpha(\cdot,\cdot) \in L^2(M \times M)$ with a vector in $\mathfrak{h} \otimes \mathfrak{h}$. Hence,

$$
\begin{aligned}
P(\alpha) &= \int |\alpha(x,y)|^2 \, V(x,y)\, dx\, dy = \langle \alpha |\, V\, \alpha \rangle \\
&\leq a\, \langle \alpha |\, \{(h+b) \otimes \mathbf{1} + \mathbf{1} \otimes (h+b)\}\, \alpha \rangle \qquad \text{(II.50)}\\
&= a \,\mathrm{Tr}\big\{(h+b)(\alpha^*\alpha + \alpha\alpha^*)\big\} = 2a \,\mathrm{Tr}\big\{(h+b)\alpha\alpha^*\big\},
\end{aligned}
$$

additionally using the reality of $\mathrm{Tr}\{(h+b)\alpha\alpha^*\}$ and $\alpha^T = -\alpha$. Similarly we obtain

$$
P(\gamma) = \langle \gamma |\, V\, \gamma \rangle \leq 2a \,\mathrm{Tr}\big\{(h+b)\gamma^2\big\}. \qquad \text{(II.51)}
$$

Eq (II.47) follows from adding up (II.50) and (II.51) and using $\gamma^2 + \alpha\alpha^* \leq \gamma$.

To prove (II.46) and later (II.49), we introduce $\vartheta := \sqrt{\gamma}$ and $\vartheta_z(x) := \vartheta(z, x)$. So, for almost all $x, x' \in M$, we have that

$$\gamma(x,x') \; = \; \int \overline{\vartheta_z(x)} \vartheta_z(x') \, dz, \qquad \rho_\gamma(x) \; = \; \int |\vartheta_z(x)|^2 \, dz, \tag{II.52}$$

These identities imply that

$$\begin{aligned} D(\rho_\gamma, \rho_\gamma) \; &= \; \int |\vartheta_w(x)|^2 \, |\vartheta_z(y)|^2 \, V(x,y) \, dx \, dy \, dw \, dz \\ &= \; \int \langle \vartheta_w \otimes \vartheta_z | \, V \, (\vartheta_w \otimes \vartheta_z) \rangle \, dw \, dz, \qquad \text{(II.53)} \\ P(\gamma) \; &= \; \int \vartheta_w(x) \, \overline{\vartheta_w(y)} \, \overline{\vartheta_z(x)} \, \vartheta_z(y) \, V(x,y) \, dx \, dy \, dw \, dz \\ &= \; \int \langle \vartheta_z \otimes \vartheta_w | \, V \, (\vartheta_w \otimes \vartheta_z) \rangle \, dw \, dz. \qquad \text{(II.54)} \end{aligned}$$

By means of (II.53), we obtain (II.46) as follows,

$$\begin{aligned} D(\rho_\gamma, \rho_\gamma) \; &= \; \int \langle \vartheta_w \otimes \vartheta_z | \, V \, (\vartheta_w \otimes \vartheta_z) \rangle \, dw \, dz \\ &\leq \; a \int \langle \vartheta_w \otimes \vartheta_z | \, \{ (h+b) \otimes \mathbf{1} + \mathbf{1} \otimes (h+b) \} \, (\vartheta_w \otimes \vartheta_z) \rangle \, dw \, dz \\ &= \; 2a \Big(\int \langle \vartheta_w | (h+b) \vartheta_w \rangle \, dw \Big) \Big(\int \langle \vartheta_z | \vartheta_z \rangle \, dz \Big) \\ &= \; 2a \, N \, \mathrm{Tr}\{ (h+b) \gamma \}. \qquad \text{(II.55)} \end{aligned}$$

(ii) Since $V \geq 0$ is nonnnegative, the Cauchy-Schwarz Inequality implies

$$|\langle \vartheta_w \otimes \vartheta_z | \, V \, (\vartheta_z \otimes \vartheta_w) \rangle| \; \leq \; \langle \vartheta_w \otimes \vartheta_z | \, V \, (\vartheta_w \otimes \vartheta_z) \rangle, \tag{II.56}$$

which, integrated over w and z, immediately yields (II.49). The semiboundedness and the coercivity of $\mathcal{E}_{BHF}^+$ is a consequence of this and $P(\alpha) \geq 0$. We remark that $P(\alpha)$ is not bounded by $D(\rho_\gamma, \rho_\gamma)$, in general.

(iii) is an immediate consequence of (II.46) and (II.47). □

Next we discuss weak lower semicontinuity. This property is of key importance for proving existence of minimizers in the calculus of variations. It turns out, however, that for attractive systems weak lower semicontinuity does not hold [11]. For example, if $V(x,y) := |x-y|^{-1}$ is the Newtonian gravitational potential, $\gamma, \tilde{\gamma}$ are two 1-pdm on $\mathfrak{h} := L^2(\mathbb{R}^3)$, and τ_λ denotes translation by $\lambda e_1 \in \mathbb{R}^3$, i.e., $[\tau_\lambda \varphi](x) := \varphi(x - \lambda e_1)$, then

$$\eta_\lambda \; := \; \gamma \; + \; \tau_\lambda \, \tilde{\gamma} \, \tau_\lambda^* \quad \longrightarrow \quad \gamma, \tag{II.57}$$

weakly in $\mathcal{L}^2(\mathfrak{h})$, as $\lambda \to \infty$, but

$$D(\eta_\lambda, \eta_\lambda) \quad \to \quad D(\gamma, \gamma) + D(\tilde{\gamma}, \tilde{\gamma}). \tag{II.58}$$

Yet, $\mathcal{E}^{-}_{BHF}$ does have a minimizer, and this is essentially due to

$$D(\eta_\lambda, \eta_\lambda) \; > \; D(\gamma, \gamma) \; + \; D(\tilde{\gamma}, \tilde{\gamma}), \tag{II.59}$$

for all $\lambda \in \mathbb{R}$, which shows that minimization tends to localize the spatial support of the 1-pdm. By a variant of Lion's concentration compactness, this idea is implemented in [11] and yields the existence of a minimizer.

Minimizers of energy functionals are necessarily stationary points of these and fulfill the corresponding Euler-Lagrange equations. In the context of HF or BHF theory, these are also often called *self-consistent equations*. We formulate these only for HF theory because its later use in (III.159). We assume that $V \geq 0$ is relative h-form bounded with relative bound $a < \infty$ in the repulsive case. In the attractive case, we require $a < (N+1)^{-1}$, but we keep in mind that we may weaken this assumption to $a < cN^{-2/3}$ according to the remark made after Lemma 1. Then, for a given 1-pdm $0 \leq \gamma \leq 1$, with $\mathrm{Tr}\, \gamma = N$ and $\mathrm{Tr}\{h\gamma\} < \infty$, we introduce the effective Hamiltonian $H(\gamma)$ by

$$\big[H(\gamma)\varphi\big](x) \;=\; [h\varphi](x) \pm \int \big\{\gamma(y,y)\, \varphi(x) - \gamma(x,y)\, \varphi(y)\big\} W(x,y)\, dy, \tag{II.60}$$

where $\varphi \in \mathfrak{d}$ is assumed to be normalized, and $\pm$ denote the repulsive (+) and attractive (-) cases, respectively.

Lemma 2 *Suppose that $V \geq 0$ is relative h-form bounded with relative bound $a < \infty$ in the repulsive case (+) and $a < N^{-1}$ in the attractive case (-), where $N > 0$. Then, for a given 1-pdm $0 \leq \gamma \leq 1$, with $\mathrm{Tr}\, \gamma = N$ and $\mathrm{Tr}\{h\gamma\} < \infty$, the effective Hamiltonian $H(\gamma)$ defines a semibounded quadratic form.*

Proof: We introduce $\vartheta := \sqrt{\gamma}$ and $\vartheta_z(x) := \vartheta(z,x)$, as in the proof of Lemma 1. Then $\langle \varphi |\, H(\gamma)\varphi \rangle$ can be written as

$$\langle \varphi |\; H(\gamma)\varphi \rangle \;=\; \langle \varphi | h\varphi \rangle \pm \int \Big\{ \langle \varphi \otimes \vartheta_z |\, V\, (\varphi \otimes \vartheta_z) \rangle - \langle \vartheta_z \otimes \varphi |\, V\, (\varphi \otimes \vartheta_z) \rangle \Big\}\, dz. \tag{II.61}$$

By the Cauchy-Schwarz Inequality,

$$\langle \varphi \otimes \vartheta_z |\, V\, (\varphi \otimes \vartheta_z) \rangle \;\geq\; \langle \vartheta_z \otimes \varphi |\, V\, (\varphi \otimes \vartheta_z) \rangle, \tag{II.62}$$

and we therefore have

$$\langle \varphi |\, H(\gamma)\varphi \rangle \;\geq\; \langle \varphi |\, h\varphi \rangle \;\geq\; e_0, \tag{II.63}$$

in the repulsive case (+).

In the attractive case (-), we observe that

$$\begin{aligned}
&\langle\varphi|\, H(\gamma)\varphi\rangle && \text{(II.64)}\\
&\geq\ \langle\varphi|\, h\varphi\rangle\ -\ \int \langle \varphi\otimes\vartheta_z|\, V\,(\varphi\otimes\vartheta_z)\rangle\, dz \\
&\geq\ \langle\varphi|\, h\varphi\rangle\ -\ a\int \Big\langle \varphi\otimes\vartheta_z\Big|\, \{(h+b)\otimes\mathbf{1}+\mathbf{1}\otimes(h+b)\}\,(\varphi\otimes\vartheta_z)\Big\rangle\, dz \\
&\geq\ -b\ +\ (1-aN)\,\langle\varphi|\,(h+b)\varphi\rangle\ -\ a\ \mathrm{Tr}\{(h+b)\gamma\} \\
&\geq\ -b\ -\ a\ \mathrm{Tr}\{(h+b)\gamma\}. && \square
\end{aligned}$$

We remark that much of the conclusion of Lieb's variational principle (I.8) holds true for general potentials (repulsive, attractive, or even indefinite) under the additional assumption of existence of a minimizer, i.e., that the infimum on the right side of (I.8) is actually a minimum. We prove this assertion in Lemma 3 below.

Lemma 3 *Let $N \in \mathbb{N}$ and γ_0 be a minimizer of $\mathcal{E}_{HF}$ with $\mathrm{Tr}(\gamma_0) = N$. Then the effective Hamiltonian $H(\gamma_0)$ has at least N eigenvalues (counting multiplicities) below its continuous spectrum, and γ_0 necessarily fulfills the self-consistent equation*

$$\mathbf{1}_{N-1}\big[H(\gamma_0)\big]\ \leq\ \gamma_0\ \leq\ \mathbf{1}_N\big[H(\gamma_0)\big], \tag{II.65}$$

where $\mathbf{1}_k[H(\gamma_0)]$ denotes the projection onto the k lowest eigenvalues $H(\gamma_0)$.

Proof: Let $0 \leq \gamma_0 \leq 1$ be a minimizer and $0 \leq \gamma_1 \leq 1$ be some other 1-pdm of trace $\mathrm{Tr}\{\gamma_1\} = \mathrm{Tr}\{\gamma_0\} = N$. Set $\gamma_\lambda := (1-\lambda)\gamma_0 + \lambda\gamma_1$, so $0 \leq \gamma_\lambda \leq 1$ and $\mathrm{Tr}\{\gamma_\lambda\} = N$, for all $0 \leq \lambda \leq 1$. An easy computation yields

$$\mathcal{E}_{HF}(\gamma_\lambda) - \mathcal{E}_{HF}(\gamma_0)\ =\ \lambda\ \mathrm{Tr}\big\{H_0\,(\gamma_1-\gamma_0)\big\}\ +\ \lambda^2\, Q[\gamma_1-\gamma_0,\ \gamma_1-\gamma_0], \tag{II.66}$$

where we abbreviate $H_0 := H(\gamma_0)$ and use

$$Q[\eta,\kappa]\ :=\ \int \big\{\eta(x,x)\,\kappa(y,y)\ -\ \eta(x,y)\,\kappa(y,x)\big\} W(x,y)\, dx\, dy. \tag{II.67}$$

Since γ_0 is a minimizer, this yields

$$\mathrm{Tr}\big\{H_0\,\gamma_0\big\}\ \leq\ \mathrm{Tr}\big\{H_0\,\gamma_1\big\} \tag{II.68}$$

by taking the limit $\lambda \to 0$. Now let $\Sigma := \inf\sigma_{\mathrm{cont}}[H_0]$ and $e_1 \leq e_2 \leq \cdots \leq e_M < \Sigma$ be the eigenvalues of H_0 (counting multiplicities) strictly below Σ (possibly, $M = 0$ or $M = \infty$).

First assume that $M < N$. Since $e_M < \Sigma$ and $\mathrm{Tr}\{\gamma_0\} = N > M$, we have

$$\delta \ := \ \mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_0\Big\} \ - \ \sum_{m=1}^{M}(e_m - \Sigma) \ > \ 0. \tag{II.69}$$

By the definition of the continuous spectrum, γ_1 can be chosen as $\mathbf{1}_M[H_0]+\gamma_1'$, where γ_1' is a rank $N-M$ projection with $\gamma_1'\,\mathbf{1}_M[H_0] = \mathbf{1}_M[H_0]\,\gamma_1' = 0$ and $0 \leq \mathrm{Tr}\{(H_0-\Sigma)\,\gamma_1'\} \leq \delta/2$. Thus

$$\begin{aligned}
&\mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_0\Big\} \ - \ \mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_1\Big\} \\
&= \ \mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_0\Big\} \ - \ \sum_{n=1}^{M}(e_n - \Sigma) \ - \ \mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_1'\Big\} \\
&\geq \ \frac{\delta}{2} \ > \ 0,
\end{aligned} \tag{II.70}$$

which implies that

$$\begin{aligned}
\mathrm{Tr}\Big\{H_0\,\gamma_0\Big\} \ &= \ \mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_0\Big\} + \Sigma\,N \\
&> \ \mathrm{Tr}\Big\{(H_0 - \Sigma)\,\gamma_1\Big\} + \Sigma\,N \ = \ \mathrm{Tr}\Big\{H_0\,\gamma_1\Big\},
\end{aligned} \tag{II.71}$$

contradicting (II.68). Therefore, $M \geq N$ and $e_N < \Sigma$, which, in particular, implies

$$\mathrm{Tr}\Big\{H_0\,\gamma_0\Big\} \ \geq \ \sum_{n=1}^{N} e_n. \tag{II.72}$$

Note that Inequality (II.72) is strict unless

$$\gamma_0 \ = \ \mathbf{1}[H_0 < e_N] + \tilde{\gamma}_0, \qquad \mathrm{Ran}\,\tilde{\gamma}_0 \ \subseteq \ \mathrm{Ran}\,\mathbf{1}[H_0 = e_N], \tag{II.73}$$

for which it becomes an equality. $\square$

II.1 The Bogolubov-Hartree-Fock Approximation for Repulsive Potentials

We begin this section by defining the Bogolubov-Hartree-Fock (BHF) functional

$$\begin{aligned}
&\mathcal{E}_{BHF}(\Gamma) \ := \\
&\quad \mathrm{Tr}\{h\,\gamma\} \ + \ \frac{1}{2}\int \Big\{\rho_\gamma(x)\rho_\gamma(y) - |\gamma(x,y)|^2 + |\alpha(x,y)|^2\Big\}\, W(x,y)\,dx\,dy,
\end{aligned} \tag{II.74}$$

with $\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A}$, such that all terms on the right side exist, and then the BHF energy by

$$E_{BHF}(N) \;:=\; \inf\big\{\mathcal{E}_{BHF}(\Gamma) \;\big|\; \Gamma \in \mathfrak{A}_N\big\}. \tag{II.75}$$

Note that, due to (I.36), we have that

$$\omega(\mathbb{H}) \;=\; \mathcal{E}_{BHF}(\Gamma_\omega), \tag{II.76}$$

for any quasifree state $\omega \in \mathfrak{Q}$, and hence

$$E_{BHF}(N) \;=\; \inf\big\{\omega(\mathbb{H}) \;\big|\; \omega \in \mathfrak{Q}_N\big\}. \tag{II.77}$$

If W is repulsive, i.e., $W(x, y) \geq 0$, for all $x, y \in M$, then we observe that

$$\begin{aligned} \mathcal{E}_{BHF}\big(\Gamma(\gamma, \alpha)\big) \;&=\; \mathcal{E}_{BHF}\big(\Gamma(\gamma, 0)\big) \;+\; \frac{1}{2}\int |\alpha(x, y)|^2\, W(x, y)\, dx\, dy \\ &\geq\; \mathcal{E}_{BHF}\big(\Gamma(\gamma, 0)\big) \\ &=\; \mathcal{E}_{HF}(\gamma), \end{aligned} \tag{II.78}$$

which implies that

$$E_{BHF}(N) \;=\; \inf\big\{\mathcal{E}_{BHF}(\Gamma) \;\big|\; \Gamma \in \mathfrak{A}_N^{(0)}\big\} \;=\; E_{HF}(N). \tag{II.79}$$

In other words, *BHF energy and HF energy agree for repulsive potentials*, and the variation over more general quasifree states than the gauge invariant ones is obsolete.

III The BHF Approximation for Attractive Potentials

The main purpose of this paper is to discuss the BHF approximation in case of *attractive* pair potentials. It turns out that our result requires some additional assumptions.

Besides the attractiveness of the potential, a second important assumption is that the fermions in the system under consideration are spin-$\frac{1}{2}$ particles (e.g., neutrons). That is, we assume the one-particle Hilbert space $\mathfrak{h} = \hat{\mathfrak{h}} \otimes \mathbb{C}^2$ of a spin-$\frac{1}{2}$, where $\hat{\mathfrak{h}}$ is again assumed to be a Hilbert space $L^2(M)$ of square-integrable functions on a configuration space M. Under this assumption we essentially prove that the minimizing state is a spin singlet state, i.e., it allows us to move the antisymmetry requirement of the variational state under consideration entirely to its spin part; see Eqs. (III.88) and (III.97).

Even though we assume the fermions to have spin $\frac{1}{2}$, both the one-body Hamiltonian $h : \mathfrak{h} \to \mathfrak{h}$ and the two-body interaction $W = -V : \mathfrak{h} \otimes \mathfrak{h} \to \mathfrak{h} \otimes \mathfrak{h}$ (note the minus sign, $V > 0$) are supposed to be diagonal in spin space and real. More precisely, we require that $h := \hat{h} \otimes \mathbf{1}_{\mathbb{C}^2}$, where $\hat{h} : \hat{\mathfrak{h}} \to \hat{\mathfrak{h}}$ is a real quadratic form defined on a dense domain $\hat{\mathfrak{d}} \times \hat{\mathfrak{d}} \subseteq \hat{\mathfrak{h}} \times \hat{\mathfrak{h}}$ and bounded below by

$$e_0 \ := \ \inf \big\{ \langle \varphi | \, \hat{h} \, \varphi \rangle \ \big| \ \varphi \in \hat{\mathfrak{d}}, \ \ \|\varphi\| = 1 \big\} \ > \ -\infty. \tag{III.80}$$

Similarly, we assume that there is a family $\{\widehat{X}_\eta\}_{\eta \in \mathcal{N}} \subseteq \mathcal{B}(\hat{\mathfrak{h}})$ of real, orthogonal projections and a positive measure μ on a measure space $\mathcal{N}$, such that

$$V \ = \ \int_{\mathcal{N}} X_\eta \otimes X_\eta \, d\mu(\eta). \tag{III.81}$$

where $X_\eta := \widehat{X}_\eta \otimes \mathbf{1}_{\mathbb{C}^2}$. The class of interaction potentials specified by (III.81) contains relevant examples, such as the Newtonian potential

$$\frac{1}{|x-y|} \ = \ \int_0^\infty \frac{dr}{\pi \, r^5} \int_{\mathbb{R}^3} d^3 z \ \mathbf{1}_{B(z,r)}(x) \ \mathbf{1}_{B(z,r)}(y), \tag{III.82}$$

for neutrons interacting by gravitational forces, where $B(z,r) := \{x \in \mathbb{R}^3 : |x - z < r\}$ denotes the ball or radius r about z, or

$$\delta_{x,y} \ = \ \sum_{z \in M} \delta_{x,z} \, \delta_{y,z}, \tag{III.83}$$

for the case of the attractive Hubbard model, with $M \subseteq \mathbb{Z}^d$ and $\delta_{x,z}$ denoting the usual Kronecker delta. It is easy to see that if V is a translation invariant potential on $\mathbb{R}^d$ (with trivial action on the spin variables) then the form

(III.81) implies that both V and its Fourier transform are positive. A fairly general class of potentials admitting a decomposition of the form (III.81) was identified in [10], the form (III.82) originates in [5].

III.1 Elimination of α as a Degree of Freedom

Given a g1-pdm $\Gamma = \Gamma(\gamma, \alpha) \in \mathfrak{A}$ with $\gamma = \gamma^*$ and $\alpha = -\alpha^T$, we define

$$\gamma \; =: \; \begin{pmatrix} \hat{\gamma}_\uparrow & \hat{\gamma}_o \\ \hat{\gamma}_o^* & \hat{\gamma}_\downarrow \end{pmatrix}, \quad \alpha \; =: \; \begin{pmatrix} \hat{\alpha}_\uparrow & \hat{\alpha}_o \\ -\hat{\alpha}_o^T & \hat{\alpha}_\downarrow \end{pmatrix}, \tag{III.84}$$

with

$$\langle f |\, \hat{\gamma}_\uparrow\, g \rangle := \langle f \otimes \delta_\uparrow |\, \gamma\, (g \otimes \delta_\uparrow) \rangle, \quad \langle f |\, \hat{\alpha}_\uparrow\, g \rangle := \langle f \otimes \delta_\uparrow |\, \hat{\alpha}\, (g \otimes \delta_\uparrow) \rangle, \tag{III.85}$$
$$\langle f |\, \hat{\gamma}_\downarrow\, g \rangle := \langle f \otimes \delta_\downarrow |\, \gamma\, (g \otimes \delta_\downarrow) \rangle, \quad \langle f |\, \hat{\alpha}_\downarrow\, g \rangle := \langle f \otimes \delta_\downarrow |\, \hat{\alpha}\, (g \otimes \delta_\downarrow) \rangle, \tag{III.86}$$
$$\langle f |\, \hat{\gamma}_o\, g \rangle := \langle f \otimes \delta_\downarrow |\, \gamma\, (g \otimes \delta_\uparrow) \rangle, \quad \langle f |\, \hat{\alpha}_o\, g \rangle := \langle f \otimes \delta_\downarrow |\, \hat{\alpha}\, (g \otimes \delta_\uparrow) \rangle, \tag{III.87}$$

where $f, g \in \hat{\mathfrak{h}}$ and $A^T(x,y) := A(y,x)$, for $x, y \in M$. Note that, apart from the antisymmetry $\hat{\alpha}_\uparrow = -\hat{\alpha}_\uparrow^T$ and $\hat{\alpha}_\downarrow = -\hat{\alpha}_\downarrow^T$, i.e.,

$$\hat{\alpha}_\uparrow(x,y) \; = \; -\hat{\alpha}_\uparrow(y,x) \quad \text{and} \quad \hat{\alpha}_\downarrow(x,y) \; = \; -\hat{\alpha}_\downarrow(y,x), \tag{III.88}$$

and the selfadjointness of $\hat{\gamma}_\uparrow$ and $\hat{\gamma}_\downarrow$, no other constraint on the matrix entries of γ and α arises from the CAR.

Inserting (III.81) into (II.74), we obtain

$$\begin{aligned} \mathcal{E}_{BHF}\big(\Gamma(\gamma,\alpha)\big) \; &:= \; \omega(H) \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(III.89)} \\ = \; \mathrm{Tr}\{h\gamma\} - \frac{1}{2} \int_{\mathcal{N}} &\Big[\big(\mathrm{Tr}\{X_\eta \gamma\} \big)^2 \; - \; \mathrm{Tr}\{X_\eta \gamma X_\eta \gamma\} \; + \; \mathrm{Tr}\{X_\eta \alpha X_\eta \alpha^*\} \Big] d\mu(\eta), \end{aligned}$$

and we recall the definition of the BHF energy with particle number N,

$$E_{BHF}(N) \; = \; \inf \big\{ \mathcal{E}_{BHF}(\Gamma) \; \big| \; \Gamma = \Gamma(\gamma,\alpha) \in \mathfrak{A}_N, \; \mathrm{Tr}\{h\gamma\} < \infty \big\}. \tag{III.90}$$

Our first main result is the following simplification of the variational principle (III.90).

Theorem 4 *Let $\hat{h} = \hat{h}^* = \hat{h}^T$ be a self-adjoint, semibounded operator, defined on a dense domain $\hat{\mathfrak{d}} \subseteq \hat{\mathfrak{h}}$. Assume* (III.81) *and that $\widehat{X}_\eta = \widehat{X}_\eta^2 = \widehat{X}_\eta^* = \widehat{X}_\eta^T$ are selfadjoint and real, for all $\eta \in \mathcal{N}$, such that V is defined on $\mathfrak{d} \otimes \mathfrak{d}$ and is an relatively bounded perturbation of $h \otimes \mathbf{1} + \mathbf{1} \otimes h$. Then, for all $N \in \mathbb{N}$,*

$$E_{BHF}(N) \; = \; 2\, \widehat{E}_{BHF}(N/2) \; := \; \inf \big\{ 2 \widehat{\mathcal{E}}_{BHF}(\hat{\gamma}) \; \big| \; \hat{\gamma} \in \widehat{\mathfrak{A}}_{N/2}, \; \mathrm{Tr}\{\hat{h}\hat{\gamma}\} < \infty \big\}, \tag{III.91}$$

(possibly, $E_{BHF}(N) = \widehat{E}_{BHF}(N/2) = -\infty$) where

$$\widehat{\mathfrak{A}}_{N/2} \; := \; \left\{ \hat{\gamma} \in \mathcal{B}(\hat{\mathfrak{h}}) \; \middle| \; 0 \le \hat{\gamma} = \hat{\gamma}^* = \hat{\gamma}^T \le \mathbf{1}, \; \mathrm{Tr}(\hat{\gamma}) = N/2 \right\} \tag{III.92}$$

and

$$\widehat{\mathcal{E}}_{BHF}(\hat{\gamma}) \; := \; \mathrm{Tr}\{h\hat{\gamma}\} \; - \; \mathcal{D}(\hat{\gamma}, \hat{\gamma}) \; + \; \frac{1}{2}\mathcal{P}(\hat{\gamma}, \hat{\gamma}) \; - \; \frac{1}{2}\mathcal{P}\big(\sqrt{\hat{\gamma} - \hat{\gamma}^2}, \sqrt{\hat{\gamma} - \hat{\gamma}^2}\big), \tag{III.93}$$

and with

$$\mathcal{D}(a,b) \; := \; \int_{\mathcal{N}} \mathrm{Tr}\{\widehat{X}_\eta \, a\} \; \mathrm{Tr}\{\widehat{X}_\eta \, b\} \, d\mu(\eta), \tag{III.94}$$

$$\mathcal{P}(a,b) \; := \; \int_{\mathcal{N}} \mathrm{Tr}\{\widehat{X}_\eta \, a \, \widehat{X}_\eta \, b\} \, d\mu(\eta). \tag{III.95}$$

Proof: Given $\hat{\gamma} \in \widehat{\mathfrak{A}}_{N/2}$, we define $\gamma, \alpha \in \mathcal{B}(\mathfrak{h})$ by

$$\gamma \; := \; \begin{pmatrix} \hat{\gamma} & 0 \\ 0 & \hat{\gamma} \end{pmatrix} \; = \; \hat{\gamma} \otimes \mathbf{1}_{\mathbb{C}^2}, \tag{III.96}$$

$$\alpha \; := \; \begin{pmatrix} 0 & \sqrt{\hat{\gamma} - \hat{\gamma}^2} \\ -\sqrt{\hat{\gamma} - \hat{\gamma}^2} & 0 \end{pmatrix} \; = \; \sqrt{\hat{\gamma} - \hat{\gamma}^2} \otimes \tau, \tag{III.97}$$

where $\tau := \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$. Trivially, $\mathrm{Tr}\{h\gamma\} < \infty$ iff $\mathrm{Tr}\{\hat{h}\hat{\gamma}\} < \infty$. Since $\hat{\gamma}$ obeys the admissibility conditions $0 \le \hat{\gamma} = \hat{\gamma}^* = \hat{\gamma}^T \le \mathbf{1}$ and $\mathrm{Tr}(\hat{\gamma}) = N/2$, it follows that $\gamma = \gamma^* = \gamma^T$, $\alpha^T = -\alpha$, $\mathrm{Tr}(\gamma) = 2\,\mathrm{Tr}(\hat{\gamma}) = N$, and

$$\Gamma[\gamma, \alpha] \; = \; \begin{pmatrix} \hat{\gamma} & 0 & 0 & \sqrt{\hat{\gamma} - \hat{\gamma}^2} \\ 0 & \hat{\gamma} & -\sqrt{\hat{\gamma} - \hat{\gamma}^2} & 0 \\ 0 & -\sqrt{\hat{\gamma} - \hat{\gamma}^2} & \mathbf{1} - \hat{\gamma} & 0 \\ \sqrt{\hat{\gamma} - \hat{\gamma}^2} & 0 & 0 & \mathbf{1} - \hat{\gamma} \end{pmatrix}. \tag{III.98}$$

It is easily checked that $\Gamma[\gamma, \alpha]$ is a real, orthogonal projection and, in particular, a g1-pdm, $0 \le \Gamma[\gamma, \alpha] \le \mathbf{1}$. Moreover, $\mathrm{Tr}(\gamma) = N$, so

$$\Gamma[\gamma, \alpha] \; \in \; \mathfrak{A}_N. \tag{III.99}$$

Moreover, with γ, α according to (III.96),

$$\begin{aligned}
&\mathcal{E}_{BHF}\big(\Gamma[\gamma,\alpha]\big) \\
&\quad = \ \mathrm{Tr}\{h\gamma\} \ - \\
&\qquad \frac{1}{2}\int_{\mathcal{N}} \Big[\big(\mathrm{Tr}\{X_\eta \gamma\}\big)^2 \ - \ \mathrm{Tr}\{X_\eta \gamma X_\eta \gamma\} \ + \ \mathrm{Tr}\{X_\eta \alpha X_\eta \alpha^*\} \Big] d\mu(\eta) \\
&\quad = \ 2\,\mathrm{Tr}\{h\hat\gamma\} \ - \ \int_{\mathcal{N}} \Big[2\big(\mathrm{Tr}\{\widehat{X}_\eta \hat\gamma\}\big)^2 \ - \ \mathrm{Tr}\{\widehat{X}_\eta \hat\gamma \widehat{X}_\eta \hat\gamma\} \\
&\qquad\qquad\qquad + \ \mathrm{Tr}\,\big\{\widehat{X}_\eta \sqrt{\hat\gamma - \hat\gamma^2}\, \widehat{X}_\eta \sqrt{\hat\gamma - \hat\gamma^2}\big\} \Big] d\mu(\eta) \\
&\quad = \ 2\widehat{\mathcal{E}}_{BHF}(\hat\gamma).
\end{aligned} \tag{III.100}$$

Therefore,

$$\begin{aligned}
2\,\widehat{E}_{BHF}(N/2) \ &= \ \inf\big\{ 2\widehat{\mathcal{E}}_{BHF}(\hat\gamma) \ \big| \ \hat\gamma \in \widehat{\mathfrak{A}}_{N/2} \ \big\} \\
&= \ \inf\big\{ \mathcal{E}_{BHF}(\hat\gamma \otimes \mathbf{1}_{\mathbb{C}^2}, \sqrt{\hat\gamma - \hat\gamma^2} \otimes \tau) \ \big| \ \hat\gamma \in \widehat{\mathfrak{A}}_{N/2} \ \big\} \\
&\geq \ E_{BHF}(N).
\end{aligned} \tag{III.101}$$

Conversely, suppose that we are given $(\gamma, \alpha) \in \mathfrak{A}_N$, with

$$\gamma \ = \ \begin{pmatrix} \hat\gamma_\uparrow & \hat\gamma_o \\ \hat\gamma_o^* & \hat\gamma_\downarrow \end{pmatrix}, \quad \alpha \ = \ \begin{pmatrix} \hat\alpha_\uparrow & \hat\alpha_o \\ -\hat\alpha_o^T & \hat\alpha_\downarrow \end{pmatrix}. \tag{III.102}$$

As $(\gamma, \alpha) \in \mathfrak{A}_N$, we have $\gamma = \gamma^*$, $\alpha = -\alpha^T$, $\mathrm{Tr}(\gamma) = N$, and $0 \leq \Gamma[\gamma, \alpha] \leq \mathbf{1}$. We then define

$$\hat\gamma \ := \ \frac{1}{4}\big(\hat\gamma_\uparrow + \hat\gamma_\downarrow + \hat\gamma_\uparrow^T + \hat\gamma_\downarrow^T\big), \tag{III.103}$$

Since $0 \leq \Gamma[\gamma, \alpha] \leq \mathbf{1}$, we also have $0 \leq \hat\gamma_\uparrow, \hat\gamma_\downarrow \leq \mathbf{1}$ and thus

$$0 \ \leq \ \hat\gamma \ = \ \hat\gamma^* \ = \ \hat\gamma^T \leq \mathbf{1}. \tag{III.104}$$

Furthermore,

$$\mathrm{Tr}\{\hat\gamma\} \ = \ \frac{1}{2}\,\mathrm{Tr}\{\hat\gamma_\uparrow + \hat\gamma_\downarrow\} \ = \ \frac{1}{2}\,\mathrm{Tr}\{\gamma\} \ = \ \frac{N}{2}, \tag{III.105}$$

and thus

$$\hat\gamma \ \in \ \widehat{\mathfrak{A}}_{N/2}. \tag{III.106}$$

Our goal is to prove that

$$\mathcal{E}_{BHF}(\gamma, \alpha) \ \geq \ 2\widehat{\mathcal{E}}_{BHF}(\hat\gamma) \tag{III.107}$$

which then implies that $E_{BHF}(N) \geq 2\,\widehat{E}_{BHF}(N/2)$ and concludes the proof. To this end, we observe that the reality of $h = h^T$ and $\widehat{X}_\eta = \widehat{X}_\eta^T$ implies that

$$\begin{aligned}
\mathrm{Tr}\{h\gamma\} &= \mathrm{Tr}\{\hat{h}(\hat{\gamma}_\uparrow + \hat{\gamma}_\downarrow)\} = \frac{1}{2}\mathrm{Tr}\{h(\hat{\gamma}_\uparrow + \hat{\gamma}_\downarrow) + h^T(\hat{\gamma}_\uparrow^T + \hat{\gamma}_\downarrow^T)\} \\
&= 2\,\mathrm{Tr}\{\hat{h}\hat{\gamma}\}, && \text{(III.108)} \\
\mathrm{Tr}\{X_\eta\gamma\} &= \mathrm{Tr}\{\widehat{X}_\eta(\hat{\gamma}_\uparrow + \hat{\gamma}_\downarrow)\} = 2\,\mathrm{Tr}\{\widehat{X}_\eta\hat{\gamma}\}. && \text{(III.109)}
\end{aligned}$$

(In particular, $\mathrm{Tr}\{\hat{h}\hat{\gamma}\} < \infty$ iff $\mathrm{Tr}\{h\gamma\} < \infty$.) Hence, comparing $\mathcal{E}_{BHF}(\gamma,\alpha)$ to $2\widehat{\mathcal{E}}_{BHF}(\hat{\gamma})$, with $\hat{\gamma}$ given in (III.103), we obtain

$$\begin{aligned}
&\mathcal{E}_{BHF}(\gamma,\alpha) - 2\widehat{\mathcal{E}}_{BHF}(\hat{\gamma}) \\
&= \int_{\mathcal{N}} \Big[\mathrm{Tr}\{X_\eta\gamma X_\eta\gamma\} - 2\,\mathrm{Tr}\{\widehat{X}_\eta\hat{\gamma}\widehat{X}_\eta\hat{\gamma}\} && \text{(III.110)} \\
&\qquad + 2\,\mathrm{Tr}\{\widehat{X}_\eta\sqrt{\hat{\gamma}-\hat{\gamma}^2}\,\widehat{X}_\eta\sqrt{\hat{\gamma}-\hat{\gamma}^2}\} - \mathrm{Tr}\{X_\eta\alpha X_\eta\alpha^*\}\Big]\,d\mu(\eta).
\end{aligned}$$

Thus (III.107) follows from the two inequalities

$$\begin{aligned}
2\,\mathrm{Tr}\{\widehat{X}\hat{\gamma}\widehat{X}\hat{\gamma}\} &\leq \mathrm{Tr}\{X\gamma X\gamma\}, && \text{(III.111)} \\
\mathrm{Tr}\{X\alpha X\alpha^*\} &\leq 2\,\mathrm{Tr}\{\widehat{X}\sqrt{\hat{\gamma}-\hat{\gamma}^2}\,\widehat{X}\sqrt{\hat{\gamma}-\hat{\gamma}^2}\}, && \text{(III.112)}
\end{aligned}$$

which we assert to hold true for any real orthogonal projection $\widehat{X} = \widehat{X}^2 = \widehat{X}^* = \widehat{X}^T$, with $X := \widehat{X} \otimes \mathbf{1}_{\mathbb{C}^2}$.

To establish (III.111), we introduce $\tilde{\gamma} := \frac{1}{2}(\hat{\gamma}_\uparrow + \hat{\gamma}_\downarrow) = \tilde{\gamma}^*$, so $\hat{\gamma} = \frac{1}{2}(\tilde{\gamma} + \tilde{\gamma}^T)$ and thus

$$\begin{aligned}
2\,\mathrm{Tr}\{\widehat{X}\hat{\gamma}\widehat{X}\hat{\gamma}\} &= \frac{1}{2}\mathrm{Tr}\{\widehat{X}(\tilde{\gamma}+\tilde{\gamma}^T)\widehat{X}(\tilde{\gamma}+\tilde{\gamma}^T)\} \\
&= \mathrm{Tr}\{\widehat{X}\tilde{\gamma}\widehat{X}\tilde{\gamma}\} + \mathrm{Tr}\{\widehat{X}\tilde{\gamma}\widehat{X}\tilde{\gamma}^T\} \\
&\leq 2\,\mathrm{Tr}\{\widehat{X}\tilde{\gamma}\widehat{X}\tilde{\gamma}\}, && \text{(III.113)}
\end{aligned}$$

using the Cauchy-Schwarz inequality. Similarly,

$$\begin{aligned}
2\,\mathrm{Tr}\{\widehat{X}\tilde{\gamma}\widehat{X}\tilde{\gamma}\} &= \frac{1}{2}\mathrm{Tr}\{\widehat{X}(\hat{\gamma}_\uparrow+\hat{\gamma}_\downarrow)\widehat{X}(\hat{\gamma}_\uparrow+\hat{\gamma}_\downarrow)\} \\
&\leq \mathrm{Tr}\{\widehat{X}\hat{\gamma}_\uparrow\widehat{X}\hat{\gamma}_\uparrow\} + \mathrm{Tr}\{\widehat{X}\hat{\gamma}_\downarrow\widehat{X}\hat{\gamma}_\downarrow\}. && \text{(III.114)}
\end{aligned}$$

Eqs. (III.113) and (III.114) imply (III.111) because of the identity

$$\mathrm{Tr}\{X\gamma X\gamma\} = \mathrm{Tr}\{\widehat{X}\hat{\gamma}_\uparrow\widehat{X}\hat{\gamma}_\uparrow\} + \mathrm{Tr}\{\widehat{X}\hat{\gamma}_\downarrow\widehat{X}\hat{\gamma}_\downarrow\} + 2\,\mathrm{Tr}\{\widehat{X}\hat{\gamma}_o\widehat{X}\hat{\gamma}_o^*\} \quad \text{(III.115)}$$

and the fact that the last term on the right side of (III.115) is negative.

The key element of the proof is Estimate (III.112). We first write

$$G \ := \ \Gamma[\gamma, \alpha] - \Gamma[\gamma, \alpha]^2 \tag{III.116}$$

as a 4×4 matrix with $\mathcal{B}(\hat{\mathfrak{h}})$-valued entries, i.e.,

$$\Gamma[\gamma, \alpha] - \Gamma[\gamma, \alpha]^2 \ =: \ \begin{pmatrix} G_{11} & G_{12} & G_{13} & G_{14} \\ G_{21} & G_{22} & G_{23} & G_{24} \\ G_{31} & G_{32} & G_{33} & G_{34} \\ G_{41} & G_{42} & G_{43} & G_{44} \end{pmatrix}, \tag{III.117}$$

and note that $G \geq 0$ particularly implies

$$\hat{\gamma}_\uparrow - \hat{\gamma}_\uparrow^2 - \hat{\gamma}_o \hat{\gamma}_o^* - \hat{\alpha}_\uparrow \hat{\alpha}_\uparrow^* - \hat{\alpha}_o \hat{\alpha}_o^* \ = \ G_{11} \ \geq \ 0, \tag{III.118}$$

$$\hat{\gamma}_\downarrow - \hat{\gamma}_\downarrow^2 - \hat{\gamma}_o^* \hat{\gamma}_o - \hat{\alpha}_\downarrow \hat{\alpha}_\downarrow^* - \hat{\alpha}_o^T (\hat{\alpha}_o^*)^T \ = \ G_{22} \ \geq \ 0. \tag{III.119}$$

$$\tag{III.120}$$

We use again $\tilde{\gamma} := \frac{1}{2}(\hat{\gamma}_\uparrow + \hat{\gamma}_\downarrow) = \tilde{\gamma}^*$ and $\hat{\gamma} = \frac{1}{2}(\tilde{\gamma} + \tilde{\gamma}^T)$ and observe that

$$\hat{\gamma}^2 \ = \ \frac{1}{4}\big[\tilde{\gamma} + \tilde{\gamma}^T\big]^2 \ \leq \ \frac{1}{2}\big[\tilde{\gamma}^2 + (\tilde{\gamma}^2)^T\big], \tag{III.121}$$

$$\tilde{\gamma}^2 \ = \ \frac{1}{4}\big[\hat{\gamma}_\uparrow + \hat{\gamma}_\downarrow\big]^2 \ \leq \ \frac{1}{2}\big[\hat{\gamma}_\uparrow^2 + \hat{\gamma}_\downarrow^2\big], \tag{III.122}$$

so

$$\hat{\gamma}^2 \ \leq \ \frac{1}{4}\big[\hat{\gamma}_\uparrow^2 + \hat{\gamma}_\downarrow^2 + (\hat{\gamma}_\uparrow^2)^T + (\hat{\gamma}_\downarrow^2)^T\big]. \tag{III.123}$$

Thus, adding up (III.118) and (III.119) and using (III.123) and the antisymmetry of $\hat{\alpha}_{\uparrow,\downarrow}$, we obtain

$$\begin{aligned} \hat{\gamma} - \hat{\gamma}^2 \ &\geq \ \frac{1}{4}\big[\hat{\alpha}_\uparrow \hat{\alpha}_\uparrow^* + \hat{\alpha}_\downarrow \hat{\alpha}_\downarrow^* + \hat{\alpha}_o \hat{\alpha}_o^* + (\hat{\alpha}_o^* \hat{\alpha}_o)^T \\ &\qquad + (\hat{\alpha}_\uparrow \hat{\alpha}_\uparrow^*)^T + (\hat{\alpha}_\downarrow \hat{\alpha}_\downarrow^*)^T + (\hat{\alpha}_o \hat{\alpha}_o^*)^T + \hat{\alpha}_o^* \hat{\alpha}_o\big] \\ &= \ \frac{1}{4}\big[\hat{\alpha}_\uparrow \hat{\alpha}_\uparrow^* + \hat{\alpha}_\uparrow^* \hat{\alpha}_\uparrow + \hat{\alpha}_\downarrow \hat{\alpha}_\downarrow^* + \hat{\alpha}_\downarrow^* \hat{\alpha}_\downarrow \\ &\qquad + \hat{\alpha}_o \hat{\alpha}_o^* + \hat{\alpha}_o^* \hat{\alpha}_o + (\hat{\alpha}_o \hat{\alpha}_o^*)^T + (\hat{\alpha}_o^* \hat{\alpha}_o)^T\big]. \end{aligned} \tag{III.124}$$

Since $A \mapsto \sqrt{A}$ is operator monotone, this implies that

$$\Big(\sum_{\nu=1}^{8} A_\nu\Big)^{1/2} \ \leq \ 2\sqrt{\hat{\gamma} - \hat{\gamma}^2}, \tag{III.125}$$

where the positive selfadjoint operators $A_1, \ldots, A_8 \geq 0$ are defined by

$$A_1 := \hat{\alpha}_\uparrow \hat{\alpha}_\uparrow^*, \quad A_2 := \hat{\alpha}_\uparrow^* \hat{\alpha}_\uparrow, \quad A_3 := \hat{\alpha}_\downarrow \hat{\alpha}_\downarrow^*, \quad A_4 := \hat{\alpha}_\downarrow^* \hat{\alpha}_\downarrow, \tag{III.126}$$

$$A_5 := \hat{\alpha}_o \hat{\alpha}_o^*, \quad A_6 := \hat{\alpha}_o^* \hat{\alpha}_o, \quad A_7 := \hat{\alpha}_o^T (\hat{\alpha}_o^*)^T, \quad A_8 := (\hat{\alpha}_o^*)^T \hat{\alpha}_o^T. \tag{III.127}$$

If M is a densely defined, closed operator then the polar decomposition yields a unitary U such that $M = U|M|$ and $M^* = |M|U^*$. Here, the absolute value of an operator M is defined as usual: $|M| := \sqrt{M^*M}$. Note that $|M^*| = \sqrt{MM^*} = U|M|U^*$. Using the Cauchy-Schwarz inequality for traces, we then observe that

$$\begin{aligned}
\mathrm{Tr}\{\widehat{X}\, M\, \widehat{X}\, M^*\} \;&=\; \mathrm{Tr}\{\widehat{X}\, U\, |M|\, \widehat{X}\, |M|\, U^*\} \\
&= \mathrm{Tr}\left\{ \big(|M|^{1/2}\, U^*\, \widehat{X}\, U\, |M|^{1/2}\big)\big(|M|^{1/2}\, \widehat{X}\, |M|^{1/2}\big) \right\} \\
&\leq \Big(\mathrm{Tr}\{U^*\, \widehat{X}\, U\, |M|\, U^*\, \widehat{X}\, U\, |M|\} \cdot \mathrm{Tr}\{\widehat{X}\, |M|\, \widehat{X}\, |M|\} \Big)^{1/2} \\
&= \Big(\mathrm{Tr}\{\widehat{X}\, |M^*|\, \widehat{X}\, |M^*|\} \cdot \mathrm{Tr}\{\widehat{X}\, |M|\, \widehat{X}\, |M|\} \Big)^{1/2} \\
&\leq \frac{1}{2}\big(\mathrm{Tr}\{\widehat{X}\, |M^*|\, \widehat{X}\, |M^*|\} + \mathrm{Tr}\{\widehat{X}\, |M|\, \widehat{X}\, |M|\} \big). \qquad \text{(III.128)}
\end{aligned}$$

We derive the following upper bound on $\mathrm{Tr}\{X_\eta \alpha X_\eta \alpha^*\}$,

$$\begin{aligned}
&\mathrm{Tr}\{X\alpha X\alpha^*\} \\
&\quad= \mathrm{Tr}\{\widehat{X}\hat{\alpha}_\uparrow \widehat{X}\hat{\alpha}_\uparrow^*\} + \mathrm{Tr}\{\widehat{X}\hat{\alpha}_\downarrow \widehat{X}\hat{\alpha}_\downarrow^*\} + 2\,\mathrm{Tr}\{\widehat{X}\hat{\alpha}_o \widehat{X}\hat{\alpha}_o^*\} \\
&\quad= \mathrm{Tr}\{\widehat{X}\hat{\alpha}_\uparrow \widehat{X}\hat{\alpha}_\uparrow^*\} + \mathrm{Tr}\{\widehat{X}\hat{\alpha}_\downarrow \widehat{X}\hat{\alpha}_\downarrow^*\} + \mathrm{Tr}\{\widehat{X}\hat{\alpha}_o \widehat{X}\hat{\alpha}_o^*\} + \mathrm{Tr}\{\widehat{X}\hat{\alpha}_o^T \widehat{X}(\hat{\alpha}_o^*)^T\} \\
&\quad\leq \frac{1}{2}\sum_{\nu=1}^{8} \mathrm{Tr}\{\widehat{X}\, \sqrt{A_\nu}\, \widehat{X}\sqrt{A_\nu}\}, \qquad \text{(III.129)}
\end{aligned}$$

with $A_1, \ldots, A_8$ as in (III.126)–(III.127), by applying (III.128).

Now we use a (special case of a nontrivial) trace inequality due to Lieb [14]: The map from the positive operators into the positive numbers given by $(A, B) \mapsto \mathrm{Tr}\{\widehat{X}A^{1/2}\widehat{X}B^{1/2}\}$ is *concave*, i.e., given positive operators $A_1, A_2, B_1, B_2 \geq 0$ and a number $\lambda \in [0, 1]$, the estimate

$$\begin{aligned}
&\lambda\ \mathrm{Tr}\left\{\widehat{X}\, A_1^{1/2}\, \widehat{X}\, B_1^{1/2}\right\} \;+\; (1-\lambda)\ \mathrm{Tr}\left\{\widehat{X}\, A_2^{1/2}\, \widehat{X}\, B_2^{1/2}\right\} \qquad \text{(III.130)} \\
&\quad\leq\; \mathrm{Tr}\left\{\widehat{X}\, [\lambda A_1 + (1-\lambda)A_2]^{1/2}\, \widehat{X}\, [\lambda B_1 + (1-\lambda)B_2]^{1/2}\right\}
\end{aligned}$$

holds true. This of course extends to averages of positive operators, and hence we obtain

$$\begin{aligned}
\mathrm{Tr}\left\{X\alpha X\alpha^*\right\} \;&\leq\; 4\cdot\frac{1}{8}\sum_{\nu=1}^{8} \mathrm{Tr}\left\{\widehat{X}\ A_\nu^{1/2}\ \widehat{X}\ A_\nu^{1/2}\right\} \\
&\leq\; 4\,\mathrm{Tr}\left\{\widehat{X}\,\left(\frac{1}{8}\sum_{\nu=1}^{8} A_\nu\right)^{1/2} \widehat{X}\,\left(\frac{1}{8}\sum_{\nu=1}^{8} A_\nu\right)^{1/2}\right\} \\
&=\; \frac{1}{2}\,\mathrm{Tr}\left\{\widehat{X}\,\left(\sum_{\nu=1}^{8} A_\nu\right)^{1/2} \widehat{X}\,\left(\sum_{\nu=1}^{8} A_\nu\right)^{1/2}\right\}. \qquad \text{(III.131)}
\end{aligned}$$

Finally, if $A \geq B \geq 0$ are positive operators, then

$$\mathrm{Tr}\big\{\widehat{X}\, B\, \widehat{X}\, B\big\} \;\leq\; \mathrm{Tr}\big\{\widehat{X}\, B\, \widehat{X}\, A\big\} \;\leq\; \mathrm{Tr}\big\{\widehat{X}\, A\, \widehat{X}\, A\big\}, \tag{III.132}$$

which, together with (III.125) and (III.131) yields the desired inequality (III.112):

$$\begin{aligned} \mathrm{Tr}\{X\alpha X\alpha^*\} \;&\leq\; \frac{1}{2}\,\mathrm{Tr}\Big\{\widehat{X}\,\Big(\sum_{\nu=1}^{8} A_\nu\Big)^{1/2}\,\widehat{X}\,\Big(\sum_{\nu=1}^{8} A_\nu\Big)^{1/2}\Big\} \qquad \text{(III.133)} \\ &\leq\; 2\,\mathrm{Tr}\big\{\widehat{X}\,\sqrt{\hat\gamma-\hat\gamma^2}\,\widehat{X}\,\sqrt{\hat\gamma-\hat\gamma^2}\big\}. \qquad \square \end{aligned}$$

III.2 Occurence of Pairing, $\alpha \neq 0$

Our next goal is to derive sufficient conditions for the occurence of *pairing* $\alpha_0 \neq 0$ for minimizers $\Gamma[\gamma_0, \alpha_0] \in \mathfrak{A}_N$ of $\mathcal{E}_{BHF}$. In view of (II.75) and (II.79), this is equivalent to

$$E_{HF}(N) \;>\; E_{BHF}(N), \tag{III.134}$$

Thanks to Theorem 4, Eq. (III.134) is equivalent to proving that the minimizers $\hat\gamma_0$ of $\widetilde{\mathcal{E}}_{BHF}$ are not orthogonal projections. Put differently, pairing occurs iff

$$\widehat{E}_{HF}(N/2) \;>\; \widehat{E}_{BHF}(N/2), \tag{III.135}$$

where

$$\widehat{E}_{BHF}(N/2) \;:=\; \inf\big\{\widehat{\mathcal{E}}_{BHF}(\hat\gamma)\;\big|\;\hat\gamma\in\widehat{\mathfrak{A}}_{N/2},\ \mathrm{Tr}\{\hat h\hat\gamma\}<\infty\big\}, \tag{III.136}$$

$$\widehat{E}_{HF}(N/2) \;:=\; \inf\big\{\widehat{\mathcal{E}}_{HF}(\hat\gamma)\;\big|\;\hat\gamma\in\widehat{\mathfrak{A}}_{N/2},\ \mathrm{Tr}\{\hat h\hat\gamma\}<\infty\big\}, \tag{III.137}$$

and

$$\widehat{\mathcal{E}}_{HF}(\hat\gamma) \;:=\; \mathrm{Tr}\{h\hat\gamma\} \;-\; \mathcal{D}(\hat\gamma,\hat\gamma) \;+\; \frac{1}{2}\mathcal{P}(\hat\gamma,\hat\gamma), \tag{III.138}$$

$$\widehat{\mathcal{E}}_{BHF}(\hat\gamma) \;=\; \widehat{\mathcal{E}}_{HF}(\hat\gamma) \;-\; \frac{1}{2}\mathcal{P}\big(\sqrt{\hat\gamma-\hat\gamma^2},\sqrt{\hat\gamma-\hat\gamma^2}\big). \tag{III.139}$$

We first remark that pairing does not always occur. We demonstrate this by a simple example. For $N, K \in \mathbb{N}$, with $N < K$, and a small coupling constant $g > 0$, let $\mathfrak{h} := \mathbb{C}^K$, $V := g\mathbf{1}$, $h := \sum_{k=1}^K h_k|k\rangle\langle k|$ and $P_N := \sum_{k=1}^N |k\rangle\langle k|$, where $h_1 < h_2 < \cdots < h_K$. Our goal is to show a certain rigidity of the HF and BHF minimizer in this case: For $g > 0$ sufficiently small, P_N or $\Gamma(P_N, 0)$ persists to be the minimizer of $\mathcal{E}_{HF}$ or $\mathcal{E}_{BHF}$, respectively. Then

$$\mathcal{E}_{BHF}[\Gamma(\gamma,\alpha)] \;=\; \sum_{k=1}^K \Big\{h_k\,\rho_k - \frac{g}{2}\,|\alpha(k,k)|^2\Big\}, \tag{III.140}$$

using $\rho_k := \gamma(k,k)$ and that $\mathcal{D}(\gamma,\gamma) = \mathcal{P}(\gamma,\gamma)$ in this case. In particular, $\mathcal{E}_{BHF}[\Gamma(P_N,0)] = \sum_{k=1}^N h_k$. Since

$$|\alpha(k,k)|^2 \;\le\; |\sqrt{\gamma-\gamma^2}(k,k)|^2 \;\le\; \begin{cases} \big|\sqrt{\gamma}(k,k)\big|^2 & \le \quad \rho_k, \\ \big|\sqrt{1-\gamma}(k,k)\big|^2 & \le \; 1-\rho_k, \end{cases} \tag{III.141}$$

we have

$$\begin{aligned} & \mathcal{E}_{BHF}[\Gamma(\gamma,\alpha)] - \mathcal{E}_{BHF}[\Gamma(P_N,0)] \qquad\qquad\qquad\qquad (\text{III.142}) \\ & \ge\; -\sum_{k=1}^{N} \Big\{ \big(h_k + \frac{g}{2}\big)\,\big(1-\rho_k\big) \Big\} + \sum_{k=N+1}^{K} \Big\{ \big(h_k - \frac{g}{2}\big)\,\rho_k \Big\}, \\ & =\; \sum_{k=1}^{N} \Big\{ \big(\mu - h_k - \frac{g}{2}\big)\,\big(1-\rho_k\big) \Big\} + \sum_{k=N+1}^{K} \Big\{ \big(h_k - \mu - \frac{g}{2}\big)\,\rho_k \Big\}, \end{aligned}$$

for any $\mu \in \mathbb{R}$. Choosing $\mu := \frac{1}{2}(h_N + h_{N+1})$, we obtain

$$\forall\, g \;<\; h_{N+1} - h_N : \qquad \mathcal{E}_{BHF}[\Gamma(\gamma,\alpha)] - \mathcal{E}_{BHF}[\Gamma(P_N,0)] > 0, \tag{III.143}$$

and hence no pairing occurs.

Theorem 5 (Sufficient Condition for Pairing) *Assume the same hypothesis as in Theorem 4 and suppose that $\hat\gamma_0 \in \widehat{\mathfrak{A}}_n$ is a minimizer of $\widehat{\mathcal{E}}_{HF}$, i.e., $\widehat{\mathcal{E}}_{HF}(\hat\gamma_0) = \widehat{E}_{HF}(n)$, where $n := N/2$. Let $\widehat{H}_0 := \widehat{H}(\hat\gamma_0)$ be the corresponding effective Hamiltonian defined by*

$$\big[\widehat{H}_0\varphi\big](x) \;=\; [\hat h\varphi](x) - \int \big\{\hat\gamma_0(y,y)\,\varphi(x) - \hat\gamma_0(x,y)\,\varphi(y)\big\} V(x,y)\,dy, \tag{III.144}$$

and denote by e_k its k^{th} eigenvalue (counting multiplicities) with corresponding eigenvector $\varphi_k \in \hat{\mathfrak{h}}$. If

$$e_{n+1} - e_n \;<\; \int \Big(|\varphi_n(x)|^2\, |\varphi_n(y)|^2 \;+\; |\varphi_{n+1}(x)|^2\, |\varphi_{n+1}(y)|^2 \Big) V(x,y)\,dx\,dy \tag{III.145}$$

then pairing occurs,

$$\widehat{E}_{HF}(n) \;>\; \widehat{E}_{BHF}(n). \tag{III.146}$$

Proof: For $\hat\gamma_0, \hat\gamma_1 \in \widehat{\mathfrak{A}}_n$ we set

$$\begin{aligned} \hat\gamma_\lambda &:= (1-\lambda)\,\hat\gamma_0 \;+\; \lambda\,\hat\gamma_1 \;\in\; \widehat{\mathfrak{A}}_{N/2}, & (\text{III.147}) \\ a_\lambda &:= \hat\gamma_\lambda - \hat\gamma_\lambda^2 \;\ge\; 0, & (\text{III.148}) \\ b &:= \hat\gamma_0\,(\mathbf{1} - \hat\gamma_1)\,\hat\gamma_0 + (\mathbf{1} - \hat\gamma_0)\,\hat\gamma_1\,(\mathbf{1} - \hat\gamma_0) \;\ge\; 0, & (\text{III.149}) \end{aligned}$$

for all $0 \le \lambda \le 1$. Straightforward computation yields

$$\begin{aligned} a_\lambda &= \hat\gamma_0 + \lambda(\hat\gamma_1 - \hat\gamma_0) - \big[\hat\gamma_0 + \lambda(\hat\gamma_1 - \hat\gamma_0)\big]^2 \\ &= \hat\gamma_0 - \hat\gamma_0^2 + \lambda\big[\hat\gamma_1 - \hat\gamma_0 + 2\hat\gamma_0^2 - \hat\gamma_0\hat\gamma_1 - \hat\gamma_1\hat\gamma_0\big] - \lambda^2(\hat\gamma_1 - \hat\gamma_0)^2 \\ &= \hat\gamma_0 - \hat\gamma_0^2 + \lambda\big[\hat\gamma_1 - \hat\gamma_1^2 - \hat\gamma_0 + \hat\gamma_0^2 + (\hat\gamma_1 - \hat\gamma_0)^2\big] - \lambda^2(\hat\gamma_1 - \hat\gamma_0)^2 \\ &= (1-\lambda)\, a_0 + \lambda\, a_1 + (\lambda - \lambda^2)\,(\hat\gamma_1 - \hat\gamma_0)^2 \end{aligned} \tag{III.150}$$

and

$$\begin{aligned} (\hat\gamma_0 - \hat\gamma_1)^2 &= \hat\gamma_0^2 + \hat\gamma_1^2 - \hat\gamma_0\hat\gamma_1 - \hat\gamma_1\hat\gamma_0 \\ &= \hat\gamma_0^2 + \hat\gamma_1 - \hat\gamma_0\hat\gamma_1 - \hat\gamma_1\hat\gamma_0 - a_1 \\ &= \hat\gamma_0(\mathbf{1} - \hat\gamma_1)\hat\gamma_0 - \hat\gamma_0\hat\gamma_1(\mathbf{1} - \hat\gamma_0) + \hat\gamma_1(\mathbf{1} - \hat\gamma_0) - a_1 \\ &= b - a_1. \end{aligned} \tag{III.151}$$

Since $(\hat\gamma_0 - \hat\gamma_1)^2 \ge 0$, this implies that

$$0 \;\le\; a_1 \;\le\; b \;\le\; \mathbf{1}, \tag{III.152}$$

and by inserting (III.151) into (III.150), we obtain

$$a_\lambda \;=\; (1-\lambda)\, a_0 + (\lambda - \lambda^2)\, b - \lambda^2\, a_1. \tag{III.153}$$

Recall that $\hat\gamma_0 \in \widehat{\mathfrak{A}}_n$ is a minimizer of $\widehat{\mathcal{E}}_{HF}$, i.e., $\widehat{\mathcal{E}}_{HF}(\hat\gamma_0) = \widehat{E}_{HF}(N/2)$. If $n \notin \mathbb{N}$ then $\hat\gamma_0$ cannot possibly be an orthogonal projection, and (III.135) holds true trivially. So, we may henceforth assume that $n \in \mathbb{N}$, i.e., $N = 2n$ is an even integer and $\hat\gamma_0$ an orthogonal projection of rank $n = N/2$. In fact, according to Lemma 3,

$$\hat\gamma_0 \;=\; \mathbf{1}_n\big[\widehat{H}(\hat\gamma_0)\big], \tag{III.154}$$

is a rank-n projection onto the n lowest eigenvalues (counting multiplicities) of the effective Hamiltonian $\widehat{H}_0$ defined in (III.144). In particular, $\hat\gamma_0$ commutes with $\widehat{H}_0 := \widehat{H}(\hat\gamma_0)$. Therefore,

$$\begin{aligned} \mathrm{Tr}\{\widehat{H}_0(\hat\gamma_1 - \hat\gamma_0)\} &= \mathrm{Tr}\{(\widehat{H}_0 - e_n)(\hat\gamma_1 - \hat\gamma_0)\} \qquad \text{(III.155)} \\ &= \mathrm{Tr}\{(\widehat{H}_0 - e_n)\hat\gamma_0(\hat\gamma_1 - \hat\gamma_0)\hat\gamma_0\} + \mathrm{Tr}\{(\widehat{H}_0 - e_n)\hat\gamma_0^\perp(\hat\gamma_1 - \hat\gamma_0)\hat\gamma_0^\perp\} \\ &= \mathrm{Tr}\{-(\widehat{H}_0 - e_n)\hat\gamma_0(\mathbf{1} - \hat\gamma_1)\hat\gamma_0\} + \mathrm{Tr}\{(\widehat{H}_0 - e_n)\hat\gamma_0^\perp\,\hat\gamma_1\,\hat\gamma_0^\perp\} \\ &= \mathrm{Tr}\{|\widehat{H}_0 - e_n|\, b\}. \end{aligned}$$

Moreover, $a_0 = \hat\gamma_0 - \hat\gamma_0^2 = 0$ and therefore

$$a_\lambda \;=\; (\lambda - \lambda^2)\, b + \lambda^2\, a_1. \tag{III.156}$$

In particular,

$$(\lambda - \lambda^2)\, b \;\le\; a_\lambda \;\le\; \lambda\, b, \tag{III.157}$$

because $0 \le a_1 \le b$. By the operator monotonicity of the square root, (III.157) and (III.155) yield

$$\begin{aligned} \widehat{\mathcal{E}}_{BHF}(\hat\gamma_\lambda) - \widehat{\mathcal{E}}_{BHF}(\hat\gamma_0) \;&=\; \lambda \, \mathrm{Tr}\{\widehat{H}_0(\hat\gamma_1 - \hat\gamma_0)\} \;-\; \lambda^2 \mathcal{D}(\hat\gamma_1 - \hat\gamma_0, \hat\gamma_1 - \hat\gamma_0) \\ &\quad + \frac{\lambda^2}{2}\mathcal{P}(\hat\gamma_1 - \hat\gamma_0, \hat\gamma_1 - \hat\gamma_0) \;-\; \frac{1}{2}\mathcal{P}\big(\sqrt{a_\lambda}, \sqrt{a_\lambda}\big) \\ &=\; \lambda\Big[\mathrm{Tr}\{|\widehat{H}_0 - e_n|\, b\} \;-\; \frac{1}{2}\mathcal{P}\big(\sqrt{b}, \sqrt{b}\big)\Big] \;+\; \mathcal{O}(\lambda^2). \end{aligned} \tag{III.158}$$

As a result,

$$\partial_\lambda \widehat{\mathcal{E}}_{BHF}(\hat\gamma_\lambda)\big|_{\lambda=0} \;=\; \mathrm{Tr}\{|\widehat{H}_0 - e_n|\, b\} \;-\; \frac{1}{2}\mathcal{P}\big(\sqrt{b}, \sqrt{b}\big). \tag{III.159}$$

Next we choose $\gamma_1 := \sum_{k=1}^{n-1} |\varphi_k\rangle\langle\varphi_k| + |\varphi_{n+1}\rangle\langle\varphi_{n+1}|$ and recall that $\gamma_0 := \sum_{k=1}^{n} |\varphi_k\rangle\langle\varphi_k|$. With this choice,

$$b \;=\; |\varphi_n\rangle\langle\varphi_n| \;+\; |\varphi_{n+1}\rangle\langle\varphi_{n+1}|, \tag{III.160}$$

whose insertion into (III.159) immediately gives

$$\begin{aligned} &\partial_\lambda \widehat{\mathcal{E}}_{BHF}(\hat\gamma_\lambda)\big|_{\lambda=0} \\ &\;=\; e_{n+1} - e_n - \int \Big(|\varphi_n(x)|^2\, |\varphi_n(y)|^2 + |\varphi_{n+1}(x)|^2\, |\varphi_{n+1}(y)|^2\Big) V(x,y)\, dx\, dy \\ &\;<\; 0, \end{aligned} \tag{III.161}$$

thanks to (III.145). This implies the asserted claim. □

We close with the remark that if pairing occurs then the corresponding 1-pdm has full rank.

Lemma 6 *Assume the same hypothesis as in Theorem 4 and additionally assume that $\langle\psi|V\psi\rangle > 0$, whenever $\psi \neq 0$. Furthermore, suppose that $\hat\gamma_0 \in \widehat{\mathfrak{A}}_n$ is a minimizer of $\widehat{\mathcal{E}}_{HF}$, i.e., $\widehat{\mathcal{E}}_{HF}(\hat\gamma_0) = \widehat{E}_{HF}(n)$, where $n := N/2$. If $\hat\gamma_0$ is not a projection then*

$$\langle\varphi|\, \hat\gamma_0\, \varphi\rangle \;>\; 0, \tag{III.162}$$

for all $\varphi \in \hat{\mathfrak{h}} \setminus \{0\}$.

The proof of Lemma 6 is similar to the corresponding statement in [6] and was also given in [11].

Acknowledgments: We are grateful to E. Lenzmann and M. Lewin for various useful discussions and for making their manuscript [11] accessible to us prior to publication. V. B. thanks E. Lieb and H. Siedentop for helpful discussions and explanations of [6].

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, E. H. Lieb, and J. P. Solovej. Generalized Hartree-Fock theory and the Hubbard model. *J. Stat. Phys.* , 76:3–90, 1994.

[3] O. Bratteli and D. Robinson. *Operator Algebras and Quantum Statistical Mechanics 1.* Text and Monographs in Physics. Springer-Verlag, Berlin, 2 edition, 1987.

[4] O. Bratteli and D. Robinson. *Operator Algebras and Quantum Statistical Mechanics 2.* Text and Monographs in Physics. Springer-Verlag, Berlin, 2 edition, 1996.

[5] C. Fefferman and R. de la Llave. Relativistic stability of matter – I. *Revista Matematica Iberoamericana*, 2(1, 2):119–161, 1986.

[6] R. Frank, E.H. Lieb, R. Seiringer, and H. Siedentop. Müller's exchange-correlation energy in density-matrix-functional theory. *Phys. Rev. A*, page 052517, 2007.

[7] J. Fröhlich, L. Jonsson, and E. Lenzmann. Boson stars as solitary waves. *Commun. Math. Phys.*, 274(1):1–30, 2007.

[8] J. Fröhlich, L. Jonsson, and E. Lenzmann. Effective dynamics for boson stars. *Nonlinearity*, 20:1031–1075, 2007.

[9] J. Fröhlich and E. Lenzmann. Dynamical collapse of white dwarfs in Hartree- and Hartree-Fock theory. *Commun. Math. Phys.*, 274(3):737–750, 2007.

[10] Ch. Hainzl and R. Seiringer. General decomposition of radial functions on r^n and applications to n-body quantum systems. *Lett. Math. Phys.*, 61:75–84, 2002.

[11] E. Lenzmann and M. Lewin. Minimiziers for the Hartree-Fock-Bogoliubov theory of neutron stars. *arXiv:0809.2560*, 2008.

[12] E. Lieb and H.-T. Yau. The Chandrasekar theory of stellar collapse as the limit of quantum mechanics. *Commun. Math. Phys.*, 112(1):147–174, 1987.

[13] E. H. Lieb. Variational principle for many-fermion systems. *Phys. Rev. Lett.* , 46(7):457–459, 1981.

[14] E.H. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Adv. Math.*, 11:267–288, 1973.