A variational lower bound on the ground state of a many-body system and the squaring parametrization of density matrices

Filipp Uskov[1], **Oleg Lychkovskiy**[2]

[1]Research intern, Skolkovo Institute of Science and Technology, Moscow, Russia
[2]Research scientist, Skolkovo Institute of Science and Technology, Moscow, Russia

E-mail: `fel1992@mail.ru`

Abstract. A variational *upper* bound on the ground state energy E_{gs} of a quantum system, $E_{\mathrm{gs}} \leqslant \langle \Psi | H | \Psi \rangle$, is well-known (here H is the Hamiltonian of the system and Ψ is an arbitrary wave function). Much less known is the variational *lower* bound on the ground state of a many-body translation-invariant lattice system with the Hamiltonian $H = \sum_{i=1}^{N} h_i$, where local terms h_i can be obtained from h_1 by translations (and rotations, for lattices in two and three dimensions). This bound reads $E_{\mathrm{gs}} \geqslant N \operatorname*{tr}_{\rho \in \mathbb{S}} h_1 \rho$, where $\mathbb{S}$ is some wisely chosen set of reduced density matrices ρ. The implementation of this latter variational principle is hampered by the difficulty of parameterizing the set $\mathbb{M}$, which is a necessary prerequisite for a variational procedure. The root cause of this difficulty is the nonlinear positivity constraint $\rho > 0$ which is to be satisfied by a density matrix. The squaring parametrization of the density matrix, $\rho = \tau^2 / \operatorname{tr} \tau^2$, where τ is an arbitrary (not necessarily positive) Hermitian operator, accounts for positivity automatically. We discuss how the squaring parametrization can be utilized to find variational lower bounds on ground states of translation-invariant many-body systems. As an example, we consider the Heisenberg model of spins $1/2$ in one and two dimensions.

1. Introduction

The ground state of a many-particle system is one of the central objects for studying the physics of the condensed state. The energy of the ground state as a rule cannot be calculated exactly. In strongly correlated systems, it is also difficult to apply perturbation theory. The upper energy limit can be found using the quantum variation method. It is clear that it is desirable to supplement the upper constraint with a lower constraint. Methods for obtaining a lower limit exist [2, 3, 4, 5, 6, 7], but they are much less developed than standard variational methods. In this paper, one of such methods is developed and applied. In this paper, systems with Heisenberg interactions on two-dimensional and one-dimensional lattices are considered.
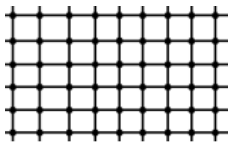
Figure 1. 2d square lattice.

Figure 2. 2d triangle lattice.

Hamiltonian of system is

$$H = J \sum_{<i,j>} (\sigma_i \sigma_j) \tag{1}$$

where $< i, j >$ is spins of neighbor atoms. Scalar and mixed products of σ-matrices denoted as

$$\begin{aligned}
(\sigma_i \sigma_j) &= \delta^{\alpha\beta} \cdot \sigma_i^\alpha \otimes \sigma_j^\beta \\
(\sigma_i \sigma_j \sigma_k) &= \varepsilon^{\alpha\beta\gamma} \cdot \sigma_i^\alpha \otimes \sigma_j^\beta \otimes \sigma_k^\gamma \\
\alpha, \beta, \gamma &\in \{x, y, z\}
\end{aligned} \tag{2}$$

A variational *upper* bound on the ground state energy E_{gs} of a quantum system, $E_{\mathrm{gs}} \leq \langle \Psi | H | \Psi \rangle$, is well-known. Challenge of this paper is to find exact lower bound on the ground state energy per 1 site in thermodynamical limit

$$E_{\mathrm{gs}} \geqslant ? \tag{3}$$

In this paper we assume that $J = 1$.

2. Main ideas
2.1. Anderson bound [3]
If we can divide full system to sum of subsystems, we can bound full ground state energy by ground state energies of subsystems:

$$H_{full} = \sum_{i=1}^{M} H_{i\ cl} \quad \Rightarrow \quad E_{gs\ full} = \sum_{i=1}^{M} E_{i\ gs\ cl} \tag{4}$$

So we tessellate full lattice by small clusters, and we get full ground state energy per 1 site $E_{gs\ full}/N$ from ground state energies of clusters $E_{gs\ cl}$:

$$E_{gs\ full}/N = \frac{d}{m} E_{gs\ cl} \tag{5}$$

where

- N - number of spins in lattice
- M - number of clusters
- d - number of bonds per 1 site
- m - number of bonds in cluster

$$M = \frac{Nd}{m} \tag{6}$$

2.2. Variational method [4, 5]
Ground state energy of full lattice can be expressed through a minimum over all density matrices $\mathbb{S}_{full}$, and it suffices to consider only those density matrices $\mathbb{S}_{full}^S$ that satisfy the same symmetries as the Hamiltonian:

$$E_{gs\ full} = \min_{\psi} \langle \psi | H | \psi \rangle = \min_{\rho_{full} \in \mathbb{S}_{full}} \mathrm{tr}\, H_{full} \rho_{full} = \min_{\rho_{full} \in \mathbb{S}_{full}^S} \mathrm{tr}\, H_{full} \rho_{full} \tag{7}$$

The crystal lattice also is divided into identical clusters, which can be used to tile the entire lattice. After that, under the trace, each expression $H_{i\ cl} \rho_{full}$ can be changed by a unitary transformation so that all $H_{i\ cl}$ match:

$$\min_{\rho_{full} \in \mathbb{S}_{full}^S} \mathrm{tr}\, H_{full} \rho_{full} = M \min_{\rho_{full} \in \mathbb{S}_{full}^S} \mathrm{tr}\, H_{cl} \rho_{full} \tag{8}$$

where M - is number of clusters in lattice. After that, the trace throughout the crystal can be divided into the sequential application of two partial traces on all the spins outside the cluster and only in the cluster:

$$M \min_{\rho_{full} \in \mathbb{S}_{full}^{S}} \operatorname{tr}_{full} H_{cl} \rho_{full} = M \min_{\rho_{full} \in \mathbb{S}_{full}^{S}} \operatorname{tr}(H_{cl} \operatorname*{tr}_{full \backslash cl} \rho_{full}) \tag{9}$$

Since the set of all cluster density matrices with symmetries of full lattice contains $\mathbb{S}_{cl}^{S} \supset \mathbb{S}_{cl}^{tr \ S} = \{\operatorname{tr}_{full \backslash cl} \rho_{full} : \rho_{full} \in \mathbb{S}_{full}^{S}\}$, so inequality arises for estimating the energy of the ground state of the crystal from below

$$E_{gs \ full} = M \min_{\rho_{cl} \in \mathbb{S}_{cl}^{tr \ S}} \operatorname{tr}_{cl} H_{cl} \rho_{cl} \geqslant M \min_{\rho_{cl} \in \mathbb{S}_{cl}^{S}} \operatorname{tr}_{cl} H_{cl} \rho_{cl} \tag{10}$$

or in another form

$$E_{gs \ full}/N \geqslant \frac{d}{m} \min_{\rho_{cl} \in \mathbb{S}_{cl}^{S}} \operatorname{tr}_{cl} H_{cl} \rho_{cl} \tag{11}$$

where we use the same notations as in (6).

2.3. Comparison

Variational method in general outperform Anderson bound for the same cluster:

$$E_{gs \ cl} = \min_{\rho_{cl} \in \mathbb{S}_{cl}} \operatorname{tr}_{cl} H_{cl} \rho_{cl} \leqslant \min_{\rho_{cl} \in \mathbb{S}_{cl}^{S}} \operatorname{tr}_{cl} H_{cl} \rho_{cl} \tag{12}$$

where $\mathbb{S}_{cl}$ is set of all density matrices of cluster.

3. Density matrix parametrization[1]

Any density matrix must satisfy 3 conditions:

$$\rho^{+} = \rho; \qquad \operatorname{tr} \rho = 1; \qquad \rho \geqslant 0 \tag{13}$$

If we take arbitrary hermitian matrix τ, and square it and normalize it, we get valid density matrix:

$$\rho = \frac{\tau^2}{\operatorname{tr} \tau^2}; \qquad \tau^{+} = \tau \tag{14}$$

If Hamiltonian is invariant with respect of some symmetry group, then τ is also invariant with respect to the same symmetry group.

In our case, the Hamiltonian is spherically symmetric; therefore, the density matrix and τ can be expressed in terms of scalar and mixed products (2). Also our Hamiltonian is invariant with respect to time inverse, but mixed products are not; therefore, density matrix and τ can be expressed in terms of scalar products:

$$\begin{aligned} \rho &= \tfrac{1}{2^n} a_i A_i; \qquad & \tau = b_i A_i \\ \{A_i\} &= \{1, \ (\sigma_j \sigma_k), \ (\sigma_j \sigma_k)(\sigma_l \sigma_m) \ , \ ...\} \end{aligned} \tag{15}$$

Parameters τ b_i run through all possible real values $a_i \in \mathbb{R}$.

Parameters ρ a_i is functions of b_i, and run through some subset $a_i = f_i(\{b\}) \in \mathbf{V} \subset \mathbb{R}$. It is convex:

$$\forall \rho_1, \rho_2 \in V, \ \forall a_1, a_2 \geq 0 \quad a_1 + a_2 = 1 \ \Rightarrow \ a_1 \rho_1 + a_2 \rho_2 \in V \tag{16}$$

Its boundary is given by the formula

$$\partial \mathbf{V} : \ \det ||\frac{\partial a_i}{\partial b_j}|| = 0 \tag{17}$$

3.1. Symbolic relations and simplifications

In Wolfram Mathematica was implemented (https://github.com/FeelUsM/ScalarMixedSpins) next operations:

Multiplication of scalar and mixed products, based on Pauli identity ($\sigma^\alpha \sigma^\beta = \delta^{\alpha\beta} + i\varepsilon^{\alpha\beta\gamma}\sigma^\gamma$), can be reduced to next formulas:

$$
\begin{aligned}
(\sigma_1\sigma_2)^2 &= 3 - 2(\sigma_1\sigma_2) \\
(\sigma_1\sigma_2)(\sigma_2\sigma_3) &= (\sigma_1\sigma_3) - i(\sigma_1\sigma_2\sigma_3) \\
(\sigma_1\sigma_2)(\sigma_1\sigma_2\sigma_3) &= -(\sigma_1\sigma_2\sigma_3) - 2i(\sigma_1\sigma_3) + 2i(\sigma_2\sigma_3) \\
(\sigma_1\sigma_2\sigma_3)(\sigma_1\sigma_2) &= -(\sigma_1\sigma_2\sigma_3) + 2i(\sigma_1\sigma_3) - 2i(\sigma_2\sigma_3) \\
(\sigma_1\sigma_2)(\sigma_2\sigma_3\sigma_4) &= (\sigma_1\sigma_3\sigma_4) - i(\sigma_1\sigma_3)(\sigma_2\sigma_4) + i(\sigma_1\sigma_4)(\sigma_2\sigma_3) \\
(\sigma_2\sigma_3\sigma_4)(\sigma_1\sigma_2) &= (\sigma_1\sigma_3\sigma_4) + i(\sigma_1\sigma_3)(\sigma_2\sigma_4) - i(\sigma_1\sigma_4)(\sigma_2\sigma_3) \\
(\sigma_1\sigma_2\sigma_3)^2 &= 6 - 2(\sigma_1\sigma_2) - 2(\sigma_1\sigma_3) - 2(\sigma_2\sigma_3) \\
(\sigma_1\sigma_2\sigma_3)(\sigma_1\sigma_2\sigma_4) &= -(\sigma_1\sigma_3)(\sigma_2\sigma_4) - (\sigma_1\sigma_4)(\sigma_2\sigma_3) + 2(\sigma_3\sigma_4) + i(\sigma_1\sigma_3\sigma_4) + i(\sigma_2\sigma_3\sigma_4) \\
(\sigma_1\sigma_2\sigma_3)(\sigma_1\sigma_4\sigma_5) &= +(\sigma_2\sigma_4)(\sigma_3\sigma_5) - (\sigma_2\sigma_5)(\sigma_3\sigma_4) - i(\sigma_1\sigma_2)(\sigma_3\sigma_4\sigma_5) + i(\sigma_1\sigma_3)(\sigma_2\sigma_4\sigma_5)
\end{aligned}
\tag{18}
$$

So if we have more than one same spin in term, we can reduce it.

In addition, in any term, we can leave only 1 mixed product:

$$
(\sigma_1\sigma_2\sigma_3)(\sigma_4\sigma_5\sigma_6) = \det \begin{pmatrix} (\sigma_1\sigma_4) & (\sigma_2\sigma_4) & (\sigma_3\sigma_4) \\ (\sigma_1\sigma_5) & (\sigma_2\sigma_5) & (\sigma_3\sigma_5) \\ (\sigma_1\sigma_6) & (\sigma_2\sigma_6) & (\sigma_3\sigma_6) \end{pmatrix}
\tag{19}
$$

We can introduce scalar product of matrices:

$$
(A, B) \equiv \operatorname{tr}(A^+ B) = \operatorname{tr}(AB); \qquad A^+ = A
\tag{20}
$$

If sets of spins in A and B differ, then $(A, B) = 0$. If A and B contains the same sets of spins, then $(A, B) = 2^N 3^C$, where N - is number of spins and C - is the number of cycles arising from the imposition of bonds formed by scalar products:

$$
\operatorname{tr}((\sigma_i\sigma_j)(\sigma_j\sigma_k)...(\sigma_l\sigma_m)(\sigma_m\sigma_i)) = 3 \cdot 2^N
\tag{21}
$$

– for one cycle.

$$
\operatorname{tr}((\sigma_i\sigma_j\sigma_k)(\sigma_i\sigma_j\sigma_n) (\sigma_k\sigma_l)...(\sigma_m\sigma_n)) = \operatorname{tr}((\sigma_i\sigma_j\sigma_k)(\sigma_i\sigma_j\sigma_k)) = 6 \cdot 2^N
\tag{22}
$$

3.2. Properties of basis

Let us append into basis $\{A_i\}$ next elements:

$$
\{B_i\} = \{(\sigma_p\sigma_r\sigma_s), \ (\sigma_p\sigma_r\sigma_s)(\sigma_j\sigma_k), \ (\sigma_p\sigma_r\sigma_s)(\sigma_j\sigma_k)(\sigma_l\sigma_m) \ , \ ...\}
$$

Basis $\{A_i\}$ or $\{A_i, B_j\}$ is not orthogonal, for example

$$
\begin{aligned}
A &= (\sigma_1\sigma_2)(\sigma_3\sigma_4) \\
B &= (\sigma_1\sigma_3)(\sigma_2\sigma_4) \\
C &= (\sigma_1\sigma_4)(\sigma_2\sigma_3) \\
g = \begin{pmatrix} (AA) & (AB) & (AC) \\ (BA) & (BB) & (BC) \\ (CA) & (CB) & (CC) \end{pmatrix} &= \begin{pmatrix} 9 & 3 & 3 \\ 3 & 9 & 3 \\ 3 & 3 & 9 \end{pmatrix} > 0
\end{aligned}
\tag{23}
$$

Table 1. Dependencies size of basis from number of spins.

N	2	3	4	5	10	15	20	30	40	50	60
$K(N)$	1	3	9	25	9495	1E+7	2E+10	6E+17	7E+25	2E+34	2E+43
4^N	16	64	256	1024	1048576	1E+9	1E+12	1E+18	1E+24	1E+30	1E+36

Also it is over-complete; however, all linear dependencies within the basis can be described by one formula:

$$+(\sigma_1\sigma_2)(\sigma_3\sigma_4\sigma_5) - (\sigma_1\sigma_3)(\sigma_2\sigma_4\sigma_5) + (\sigma_1\sigma_4)(\sigma_2\sigma_3\sigma_5) - (\sigma_1\sigma_5)(\sigma_2\sigma_3\sigma_4) = 0 \qquad (24)$$

– for elements with odd number of spins, and by the formula:

$$\det \begin{pmatrix} (\sigma_1\sigma_5) & (\sigma_2\sigma_5) & (\sigma_3\sigma_5) & (\sigma_4\sigma_5) \\ (\sigma_1\sigma_6) & (\sigma_2\sigma_6) & (\sigma_3\sigma_6) & (\sigma_4\sigma_6) \\ (\sigma_1\sigma_7) & (\sigma_2\sigma_7) & (\sigma_3\sigma_7) & (\sigma_4\sigma_7) \\ (\sigma_1\sigma_8) & (\sigma_2\sigma_8) & (\sigma_3\sigma_8) & (\sigma_4\sigma_8) \end{pmatrix} = 0 \qquad (25)$$

– for elements with even number of spins, but the last formula is a consequence (19) and (24). This statement was tested up to 10 spins. Its extrapolation to a greater number of spins is a mathematical hypothesis. By compiling a Gram matrix, one can always get rid of overflow vectors. In the future we will assume that we have prepared an uncrowded basis.

The number of elements in the basis $\{A_i\}$ without taking into account over-complete is equal to

$$K(N) = C_N^{2k}(2k-1)!! \qquad (26)$$

where C_N^{2k} - binomial coefficient, N is number of spins and k is number of considered pairs. $K(N)$ grows faster then exponentially (see table 1).

4. Examples and results

Let us consider 1d lattice and cluster with 4 spins.

$$H_{cl} = (\sigma_1\sigma_2) + (\sigma_2\sigma_3) + (\sigma_3\sigma_4) \qquad (27)$$

Basis will be

$$\{A_k\} = \{1, (\sigma_1\sigma_2), (\sigma_1\sigma_3), (\sigma_1\sigma_4), (\sigma_2\sigma_3), (\sigma_2\sigma_4), (\sigma_3\sigma_4), \\ (\sigma_1\sigma_2)(\sigma_3\sigma_4), (\sigma_1\sigma_3)(\sigma_2\sigma_4), (\sigma_1\sigma_4)(\sigma_2\sigma_3)\} \qquad (28)$$

Let apply squaring parametrization:

$$\tau = b_i A_i, \quad \rho = \frac{1}{2^4} a_i A_i, \quad \rho = \tau^2 \ \Rightarrow \ a_k = a_{k,ij} b_i b_j \qquad (29)$$

with normalization:

$$a_1 = 1 \qquad (30)$$

Lattice symmetries:

$$a_{(1,2)} = a_{(2,3)} = a_{(3,4)}; \ a_{(1,3)} = a_{(2,4)} \qquad (31)$$

Table 2. Results.

N	Anderson bound	variational method
4	-2.1547	-2.0
5	-1.92789	-1.86852
6	-1.99486	-1.86852

But Hamiltonian already has symmetries, which we can apply firstly:

$$1 \leftrightarrow 4, \ 2 \leftrightarrow 2, \ \Rightarrow b_{(1,2)} = b_{(3,4)}; \ \ b_{(1,3)} = b_{(2,4)} \ \Leftrightarrow \ a_{(1,2)} = a_{(3,4)}; \ \ a_{(1,3)} = a_{(2,4)} \tag{32}$$

Next restriction remains:

$$a_{(1,2)} = a_{(2,3)} \tag{33}$$

We can use numerical methods for calculating

$$\min \mathrm{tr}\, \mathrm{H}\rho = \min_{\mathrm{b}} \frac{\eta_{\mathrm{ij}} \mathrm{b}_{\mathrm{i}} \mathrm{b}_{\mathrm{j}}}{\mathrm{a}_{1,\mathrm{ij}} \mathrm{b}_{\mathrm{i}} \mathrm{b}_{\mathrm{j}}} \tag{34}$$

where $\eta_{ij} = Tr(A_i H A_j)$ with constraint

$$(a_{(1,2),ij} - a_{(2,3),ij}) b_i b_j = 0 \tag{35}$$

Results for clusters with 4,5,6 spins you can see in table 2.

References
[1] Il'in N, Shpagina E, Uskov F and Lychkovskiy O 2018 Squaring parametrization of constrained and unconstrained sets of quantum states *J. Phys.* A: Math. Theor. **51**.085301
[2] Tarrach R and Valentí R 1990 Exact lower bounds to the ground-state energy of spin systems: The two-dimensional S= 1/2 antiferromagnetic Heisenberg model *Phys. Rev.* B **41(13)** 9611
[3] Anderson PW 1951 Limits on the energy of the antiferromagnetic Ground State *Phys. Rev.* **83(6)** 1260
[4] Baumgratz T and Plenio MB 2012 Lower bounds for ground states of condensed matter systems *New Journal of Physics* **14(2)** 023027
[5] Mazziotti DA 2002 Variational minimization of atomic and molecular ground-state energies via the two-particle reduced density matrix *Phys.Rev.* A **65(6)** 062511
[6] Beach KSD and Sandvik AW 2006 Some formal results for the valence bond basis *Nuclear Physics* B **750(3)** 142178
[7] Mattis DC and Pan CY 1988 Ground-State Energy of Heisenberg antiferromagnet for Spins s=1/2 and s=1 in d=1 and 2 Dimensions *Phys.Rev.Lett.* **61(4)** 463