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

F. Uskov[1], **O. Lychkovskiy**[1,2]

[1] Skolkovo Institute of Science and Technology, Nobel street 3, Moscow 121205, Russia
[2] Steklov Mathematical Institute of Russian Academy of Sciences, Gubkina str., 8, Moscow 119991, Russia

E-mail: fel1992@mail.ru

Abstract. A variational *upper* bound on the ground state energy $E_{\rm gs}$ of a quantum system, $E_{\rm gs} \leqslant \langle\Psi|H|\Psi\rangle$, is well-known (here H is the Hamiltonian of the system and Ψ is an arbitrary wave function). Much less known are variational *lower* bounds on the ground state. We consider one such bound which is valid for a many-body translation-invariant lattice system. Such a lattice can be divided into clusters which are identical up to translations. The Hamiltonian of such a system can be written as $H = \sum_{i=1}^{M} H_i$, where a term H_i is supported on the i'th cluster. The bound reads $E_{\rm gs} \geqslant M \inf_{\rho_{cl} \in \mathbb{S}_{cl}^G} \mathrm{tr}_{cl}\, \rho_{cl}\, H_{cl}$, where $\mathbb{S}_{cl}^G$ is some wisely chosen set of reduced density matrices of a single cluster. The implementation of this latter variational principle can be 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, $\rho = \tau^2/\,\mathrm{tr}\,\tau^2$, 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 a one-dimensional Heisenberg antiferromagnet.

1. Introduction

The ground state of a many-particle system is one of the central objects studied in condensed matter physics. The ground state energy as a rule cannot be calculated exactly. In strongly correlated systems, it is also difficult to apply the perturbation theory. A common way to asses the ground state energy is via variational methods. An upper bound on the ground state energy, $E_{\rm gs} \leqslant \langle\Psi|H|\Psi\rangle$, is well-known. It is often desirable to supplement the latter with a *lower* bound. Methods for obtaining lower bounds on ground state energies of many-body systems exist [1, 3, 2, 4, 5], but they are much less developed than standard variational methods. In this paper we suggest one such method applicable to translation-invariant lattice systems with local interactions. The method is applied to a simple system, and its merits and prospects are discussed.

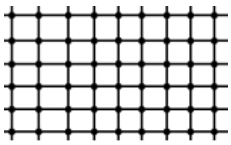

Figure 1. Square lattice.

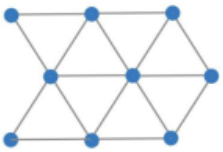

Figure 2. Triangular lattice.

2. Lower bound on the ground state energy of a translation-invariant lattice system

2.1. Our lower bound

We consider a system of spins on a lattice with N sites. The lattice is invariant with respect to the group of translation and, for two- and three-dimensional lattices, rotations. For example, this can be a linear chain in one dimension and a square or a triangular lattice in two dimensions, see Figures 1,2. Due to the symmetry, a lattice can be divided into identical clusters. The Hamiltonian of the system is defined on the lattice and is invariant under a group G which contains the symmetries of the lattice and, in general, some other symmetries,

$$U\,H\,U^\dagger = H \qquad \forall\ \ U \in G. \tag{1}$$

The Hamiltonian can be written as

$$H = \sum_{i=1}^{M} H_i, \tag{2}$$

where H_i is the local Hamiltonian of the i'th cluster, and the total number of clusters is M. In what follows we will use a special notation $H_cl \equiv H_1$ for the first cluster. Local terms H_i can be transformed one to another by group actions, i.e.

$$H_i = U\,H_j\,U^\dagger \quad \text{ for some } \quad U \in G. \tag{3}$$

Now let us derive our variational lower bound. First we note that the ground state energy of the system is given variationally by

$$E_{gs} = \inf_{\rho \in \mathbb{S}^G} \mathrm{tr}\, H\rho, \tag{4}$$

where $\mathbb{S}^G$ is a set of density matrices ρ invariant under the group G [6]. Further, using eqs. (2) and (3), we get

$$E_{gs} = M \inf_{\rho \in \mathbb{S}^G} \mathrm{tr}_{cl}\left(H_{cl}\,\mathrm{tr}_{\overline{cl}}\,\rho\right), \tag{5}$$

where tr_{cl} and $\mathrm{tr}_{\overline{cl}}$ are partial traces over the cluster and its complement, respectively. Note that while $\rho_{cl} \equiv \mathrm{tr}_{\overline{cl}}\,\rho$ is just the reduced density matrix of the cluster, variation in eq. (5) is not performed over the set of *all* ρ_{cl}. Instead, the minimization is effectively performed over those ρ_{cl} which can be obtained from $\rho \in \mathbb{S}^G$. The set of ρ_{cl} satisfying the latter condition is unknown. However, we can lower bound E_{gs} by performing minimization over a larger set $\mathbb{S}^G_{cl}$ of the reduced density matrices of the cluster symmetric under the group G. This way we obtain our variational lower bound

$$E_{gs} \geq M \inf_{\rho_{cl} \in \mathbb{S}^G_{cl}} \mathrm{tr}_{cl}\, H_{cl}\rho_{cl}. \tag{6}$$

This bound is the main general result of the present paper. It should be stressed that H_{cl} is *not* invariant under the group G, in contrast to $\rho_{cl} \in \mathbb{S}^G_{cl}$.

We further observe that this bound can be enhanced by requiring that ρ_{cl} satisfies local sum rules which follow from the anti-Hermitian Stationary Schrödinger equation [6].

2.2. Density matrix parametrization
In order to be able to perform minimization in eq. (6) one needs to parameterize the set $\mathbb{S}_{cl}^{G}$ of density matrices. Remind that any density matrix must satisfy three conditions,

$$\rho^{\dagger} = \rho; \qquad \operatorname{tr}\rho = 1; \qquad \rho > 0. \tag{7}$$

The positivity condition is nonlinear and thus the most problematic. A squaring parametrization has been developed in [7] which automatically accounts for the positivity as well as other two conditions. In addition, it is well suted for many-body systems and for accounting for symmetries. Its main idea is that if we take an arbitrary hermitian matrix τ, and square it and normalize it, we get a valid density matrix:

$$\rho = \frac{\tau^2}{\operatorname{tr}\tau^2}, \qquad \text{where} \qquad \tau^{+} = \tau. \tag{8}$$

We will us the squaring parametrization of density matrices to practically apply the bound (6).

2.3. Comparison to the Anderson bound
Let us remind the Anderson bound, which is arguably the first and the most widely used lower bound on E_{gs} [1]. It is based on a simple fact that infimum of a sum is greater than the sum of infimums. This leads to the bound

$$E_{gs} \geq M \inf_{\rho_{cl}} \operatorname{tr}_{cl} H_{cl}\rho_{cl}. \tag{9}$$

The infimum here is taken over all density matrices ρ_{cl} of a cluster. For this reason, the Anderson bound is weaker than our bound (6).

3. Application to a system of spins 1/2 with Heisenberg interactions

3.1. Spin systems with Heisenberg interactions
In the present section we show how the bound (6) can be applied to a system of N spins with the Heisenberg interaction. The Hamiltonian of this system reads

$$H = \sum_{<i,j>} (\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j), \tag{10}$$

where the sum is taken over all neighbouring sites on the lattice, $\boldsymbol{\sigma}_i$ is the vector consisting of three Pauli matrices of the i'th spin and $(\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j)$ is the corresponding scalar product of sigma-matrices. This Hamiltonian is invariant with respect to a global $SU(2)$ symmetry, in other words, to the simultaneous rotations of all spins. It is also invariant under inversion of time and respects the spatial symmetries of the lattice.

As is discussed in details in [7, 6], a density matrix ρ and matrix τ from eq. (8) can be expressed in terms of scalar products of sigma matrices,

$$\begin{aligned} &\rho = 2^{-N} \textstyle\sum_{\mathcal{A}} a_{\mathcal{A}} A_{\mathcal{A}}; \qquad\qquad \tau = \textstyle\sum_{\mathcal{A}} b_{\mathcal{A}} A_{\mathcal{A}} \\ &\{A_{\mathcal{A}}\} = \{1,\ \ (\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k),\ \ (\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k)(\boldsymbol{\sigma}_l\boldsymbol{\sigma}_m)\ \ ,\ \ ...\}, \quad i,k,l,m,... = 1,2,...N. \end{aligned} \tag{11}$$

Here b_i are arbitrary real numbers while a_i are some functions of b_i determined by eq. (8).

Let us make several remarks concerning the algebra of $A_{\mathcal{A}}$. First, we list useful relations,

$$\begin{array}{rcl}
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)^2 & = & 3-2(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3) & = & (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)-i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3) & = & -(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)-2i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)+2i(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2) & = & -(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)+2i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)-2i(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4) & = & (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)-i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4)+i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2) & = & (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)+i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4)-i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)^2 & = & 6-2(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)-2(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)-2(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4) & = & -(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4)-(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)+2(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)+i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)+i(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)\\
(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5) & = & +(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_5)-(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_5)(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)-i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5)+i(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5),
\end{array} \tag{12}$$

where $(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)$ is the mixed product of sigma matrices. Further, a product of two mixed products can always be represented through scalar products:

$$(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5\boldsymbol{\sigma}_6)=\det\begin{pmatrix}(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)\\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_5) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_5) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_5)\\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_6) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_6) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_6)\end{pmatrix}. \tag{13}$$

One can introduce scalar product on the space of operators according to

$$(A,B)\equiv \mathrm{tr}(A^+B)=\mathrm{tr}(AB);\qquad A^+=A \tag{14}$$

(not to be confused with the scalar product of sigma matrices). If supports of A and B on a lattice do not coincide, then $(A,B)=0$. If A and B have the same support, then $(A,B)=2^N3^C$, where N - is number of spins and C - is the number of cycles arising when the bonds contained in A are superimposed on the bonds contained in B (*cf.* ref. [8]). In particular,

$$\mathrm{tr}(\ (\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j)(\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k)...(\boldsymbol{\sigma}_l\boldsymbol{\sigma}_m)(\boldsymbol{\sigma}_m\boldsymbol{\sigma}_i)\ )=3\cdot 2^N \tag{15}$$

(one cycle) and

$$\mathrm{tr}(\ (\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k)(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j\boldsymbol{\sigma}_n)\ (\boldsymbol{\sigma}_k\boldsymbol{\sigma}_l)...(\boldsymbol{\sigma}_m\boldsymbol{\sigma}_n)\ )=\mathrm{tr}(\ (\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k)(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k)\ )=6\cdot 2^N. \tag{16}$$

3.2. Properties of the set $\{A_{\mathcal{A}}\}$

Let us consider a set $B_{\mathcal{B}}$ of the following form:

$$\{B_{\mathcal{B}}\}=\{(\boldsymbol{\sigma}_p\boldsymbol{\sigma}_r\boldsymbol{\sigma}_s),\ \ (\boldsymbol{\sigma}_p\boldsymbol{\sigma}_r\boldsymbol{\sigma}_s)(\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k),\ \ (\boldsymbol{\sigma}_p\boldsymbol{\sigma}_r\boldsymbol{\sigma}_s)(\boldsymbol{\sigma}_j\boldsymbol{\sigma}_k)(\boldsymbol{\sigma}_l\boldsymbol{\sigma}_m)\ \ ,\ \ ...\}$$

Bases $\{A_{\mathcal{A}}\}$ and $\{A_{\mathcal{A}},B_{\mathcal{B}}\}$ are not orthogonal, for example

$$\begin{array}{c}
\mathrm{A}=(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)\\
\mathrm{B}=(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4)\\
\mathrm{C}=(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\\
g=\begin{pmatrix}(AA) & (AB) & (AC)\\ (BA) & (BB) & (BC)\\ (CA) & (CB) & (CC)\end{pmatrix}=\begin{pmatrix}9 & 3 & 3\\ 3 & 9 & 3\\ 3 & 3 & 9\end{pmatrix}>0
\end{array} \tag{17}$$

Also they are over-complete. However, we suggest that all linear dependencies within the basis can be described by

$$+(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5)-(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5)+(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_5)-(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_5)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4)=0 \tag{18}$$

Table 1. The size $K(N)$ of the overcomplete set $\{A_{\mathcal{A}}\}$. One can see that for $N \lesssim 30$ this size is below the number of real parameters of the corresponding density matrix.

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

for elements with an odd number of spins, and by

$$\det \begin{pmatrix} (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_5) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_5) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_5) & (\boldsymbol{\sigma}_4\boldsymbol{\sigma}_5) \\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_6) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_6) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_6) & (\boldsymbol{\sigma}_4\boldsymbol{\sigma}_6) \\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_7) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_7) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_7) & (\boldsymbol{\sigma}_4\boldsymbol{\sigma}_7) \\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_8) & (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_8) & (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_8) & (\boldsymbol{\sigma}_4\boldsymbol{\sigma}_8) \end{pmatrix} = 0 \tag{19}$$

for elements with an even number of spins. Observe that but the latter formula is a consequence pof eqs. (13) and (18). We tested this hypothesis for up to 10 spins.

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

$$K(N) = \sum_{k=0}^{[N/2]} C_N^{2k}(2k-1)!! \tag{20}$$

where C_N^{2k} - binomial coefficient, N is number of spins and k is number of considered pairs. $K(N)$ grows faster then exponentially with N (see Table 1), however it can be rather small for $N \sim 10$.

3.3. An example

Let us consider one-dimensional lattice with the Heisenberg Hamiltonian (10) and a cluster with 4 spins.

$$H_{cl} = (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2) + (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3) + (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4) \tag{21}$$

The basis reads

$$\begin{aligned} \{A_k\} = \{1,\ &(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2),\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3),\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4),\ (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3),\ (\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4),\ (\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4), \\ &(\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4),\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_4),\ (\boldsymbol{\sigma}_1\boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3)\}, \quad k = 0, 1, , ...9. \end{aligned} \tag{22}$$

Let apply squaring parametrization,

$$\tau = b_k A_k, \ \ \rho = 2^{-4} a_k A_k, \quad \rho = \tau^2 \ \ \Rightarrow \ \ a_k = a_{k,ij} b_i b_j, \tag{23}$$

where summation over repeating indices is implied and the normalization is imposed by

$$a_0 = 1 \tag{24}$$

Translational invariance implies

$$a_{(1,2)} = a_{(2,3)} = a_{(3,4)}; \ \ a_{(1,3)} = a_{(2,4)} \tag{25}$$

The Hamiltonian (21) of the cluster is not translationally invariant, but it posses a remaining mirror symmetry. For this reason two of the above equalities can be satisfied seamlessly:

$$b_{(1,2)} = b_{(3,4)}; \ \ b_{(1,3)} = b_{(2,4)} \ \Rightarrow \ a_{(1,2)} = a_{(3,4)}; \ \ a_{(1,3)} = a_{(2,4)}. \tag{26}$$

Table 2. Anderson bound compared to the bound (6) for a linear chain with the Heisenberg Hamiltonian (2). Given are the lower bounds on E_{gs}/N. For an infinite chain the Bethe ansatz results reads $E_{gs}/N = 1 - 4\log 2 \simeq -1.77259$.

N	Anderson bound	bound (6)
3	-2.0	-2.0
4	-2.1547	-2.0
5	-1.9279	-1.8685
6	-1.9947	-1.8685
7	-1.8908	-1.8255

The condition

$$a_{(1,2)} = a_{(2,3)} \tag{27}$$

remains and should be accounted for during optimization.

Finally we perform a numerical search for a minimum of $\mathrm{tr}\, H_{cl}\tau^2$ with the constraints $\mathrm{tr}\,\tau^2 = 1$ and $\mathrm{tr}(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\tau^2 = (\boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3)\tau^2$. The result is presented in Table 2. The results for an analogous procedure for $N = 5, 6$ are also presented in this Table. One can see that for a given size of a cluster the bound (6) outperforms the Anderson bound. The caveat here is that for a given N the calculations for bound (6) require much more resources than those for the Anderson bound. Whether the bound (6) is able to compete with the Anderson bound in practical numerical computations is a question open for future research.

Acknowledgements.
We are grateful to E. Shpagina and N. Il'in for useful discussions. The work was supported by the Russian Science Foundation under the grant No. 17-71-20158.

References

[1] Anderson PW 1951 Limits on the energy of the antiferromagnetic Ground State *Phys. Rev.* **83(6)** 1260
[2] 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
[3] Nishimori H and Ozeki Y 1989 Ground-State Long-Range Order in the Two-Dimensional XXZ Model OR Long-range order in antiferromagnetic quantum spin systems OR Long-range order in the xxz model *J. Phys. Soc. Jpn.* **58** 1027.
[4] Schwerdtfeger CA and Mazziotti DA 2009 Convex-set description of quantum phase transitions in the transverse Ising model using reduced-density-matrix theory. *The Journal of chemical physics* **130(22)** 224102
[5] Baumgratz T and Plenio MB 2012 Lower bounds for ground states of condensed matter systems *New Journal of Physics* **14(2)** 023027
[6] Shpagina E, Uskov F, Il'in N, and Lychkovskiy O 2018 Stationary Schrdinger equation with density matrices instead of wave functions, *submitted to present proceedings*
[7] 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
[8] Beach KSD and Sandvik AW 2006 Some formal results for the valence bond basis *Nuclear Physics* B **750(3)** 142178