

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

F. Uskov¹, 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_{gs} of a quantum system, $E_{\text{gs}} \leq \langle \Psi | H | \Psi \rangle$, 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 $E_{\text{gs}} \geq M \inf_{\rho_{cl} \in \mathbb{S}_{cl}^G} \text{tr}_{cl} \rho_{cl} H_{cl}$, 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, $\rho = \tau^2 / \text{tr} \tau^2$, 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 assess the ground state energy is via variational methods. An upper bound on the ground state energy, $E_{\text{gs}} \leq \langle \Psi | H | \Psi \rangle$, 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.

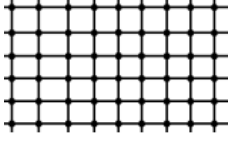


Figure 1. Square lattice.

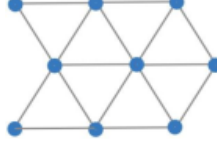


Figure 2. Triangular lattice.

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, a square or a triangular lattice in two dimensions (see figures 1, 2) *etc.* 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,

$$U H U^\dagger = H \quad \forall \quad U \in G. \quad (1)$$

The Hamiltonian can be written as

$$H = \sum_{i=1}^M H_i, \quad (2)$$

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.

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

Now let us derive our variational lower bound. It is well known that $E_{gs} = \inf_{\Psi} \langle \Psi | H | \Psi \rangle$, where infimum is taken over all normalized vectors Ψ of the Hilbert space. We observe that this equality can be alternatively formulated in terms of density matrices ρ of the closed quantum system under considerations. Remind that a density matrix is an operator which satisfies three conditions,

$$\rho^\dagger = \rho; \quad \text{tr } \rho = 1; \quad \rho > 0. \quad (4)$$

One can replace optimization over vectors Ψ by optimization over density matrices ρ :

$$E_{gs} = \inf_{\rho \in \mathbb{S}^G} \text{tr } H \rho, \quad (5)$$

Here \mathbb{S}^G is a set of density matrices ρ invariant under the group G , i.e. in addition to conditions (4) the density matrices from the set \mathbb{S}^G satisfy

$$U \rho U^\dagger = \rho \quad \text{for all } U \in G. \quad (6)$$

Indeed, the density matrix which saturates the infimum in eq. (5) is simply the normalized projection onto the ground state subspace. This observation justifies eq. (5). It should be stressed that density matrices appear here and in what follows as a formal tool for calculating a bound on the ground state of a *closed* system. Whether a system can be actually prepared in mixed states described by these density matrices is of no relevance in our argument.

Further, using eqs. (2) and (3), we get

$$E_{gs} = M \inf_{\rho \in \mathbb{S}^G} \text{tr}_{cl} (H_{cl} \text{tr}_{\overline{cl}} \rho), \quad (7)$$

where tr_{cl} and $\text{tr}_{\bar{cl}}$ are partial traces over the cluster and its complement, respectively. Note that while $\rho_{cl} \equiv \text{tr}_{\bar{cl}} \rho$ is just the reduced density matrix of the cluster, variation in eq. (7) is not performed over the set of *all* ρ_{cl} . Instead, the minimization is 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}_{cl}^G of the reduced density matrices of the cluster symmetric under the group G . This way we obtain our variational lower bound

$$E_{gs} \geq M \inf_{\rho_{cl} \in \mathbb{S}_{cl}^G} \text{tr}_{cl} H_{cl} \rho_{cl}. \quad (8)$$

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 $\rho_{cl} \in \mathbb{S}_{cl}^G$.

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. (8) one needs to parameterize the set \mathbb{S}_{cl}^G of density matrices. Among the conditions which determine this set, the positivity condition, $\rho_{cl} > 0$, is the most problematic due to its nonlinear nature. A squaring parametrization has been developed in [7] which automatically accounts for the positivity as well as other conditions of the type (4),(6). In addition, it is well-suited for many-body systems. Its main idea is that if we take an arbitrary hermitian matrix τ and square and normalize it, we get a valid density matrix:

$$\rho = \frac{\tau^2}{\text{tr} \tau^2}, \quad \text{where} \quad \tau^\dagger = \tau. \quad (9)$$

We will use the squaring parametrization of density matrices to practically apply the bound (8).

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 an infimum of a sum is greater than the sum of infima. This leads to the bound

$$E_{gs} \geq M \inf_{\rho_{cl}} \text{tr}_{cl} H_{cl} \rho_{cl}. \quad (10)$$

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 (8).

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

3.1. Spin systems with Heisenberg interactions

In the present section we demonstrate how the bound (8) can be applied to a system of N spins with the Heisenberg interaction. The Hamiltonian of this system reads

$$H = \sum_{\langle i,j \rangle} (\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j), \quad (11)$$

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 of the system (11) invariant under $SU(2)$ rotations and time inversion can be expressed in terms of scalar products of sigma matrices; the same is true for the auxiliary matrix τ from eq. (9):

$$\rho = 2^{-N} \sum_{\mathcal{A}} a_{\mathcal{A}} A_{\mathcal{A}}, \quad \tau = \sum_{\mathcal{A}} b_{\mathcal{A}} A_{\mathcal{A}}, \quad (12)$$

$$\{A_{\mathcal{A}}\} = \{1, (\sigma_j \sigma_k), (\sigma_j \sigma_k)(\sigma_l \sigma_m), \dots\}, \quad i, k, l, m, \dots = 1, 2, \dots N. \quad (13)$$

Here \mathcal{A} is a multi-index which enumerates the set (13), b_i are arbitrary real numbers while a_i are some functions of b_i determined by eq. (9).

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

Here we discuss some properties of the set $A_{\mathcal{A}}$ and a related set $B_{\mathcal{A}}$ (see below). Some of this properties have immediate consequences for the implementation of our variational lower bound, while others may prove useful in further developments.

First, we list useful relations [7]

$$\begin{aligned} (\sigma_1 \sigma_2)^2 &= 3 - 2(\sigma_1 \sigma_2) \\ (\sigma_1 \sigma_2)(\sigma_2 \sigma_3) &= (\sigma_1 \sigma_3) - i(\sigma_1 \sigma_2 \sigma_3) \\ (\sigma_1 \sigma_2)(\sigma_1 \sigma_2 \sigma_3) &= -(\sigma_1 \sigma_2 \sigma_3) - 2i(\sigma_1 \sigma_3) + 2i(\sigma_2 \sigma_3) \\ (\sigma_1 \sigma_2 \sigma_3)(\sigma_1 \sigma_2) &= -(\sigma_1 \sigma_2 \sigma_3) + 2i(\sigma_1 \sigma_3) - 2i(\sigma_2 \sigma_3) \\ (\sigma_1 \sigma_2)(\sigma_2 \sigma_3 \sigma_4) &= (\sigma_1 \sigma_3 \sigma_4) - i(\sigma_1 \sigma_3)(\sigma_2 \sigma_4) + i(\sigma_1 \sigma_4)(\sigma_2 \sigma_3) \\ (\sigma_2 \sigma_3 \sigma_4)(\sigma_1 \sigma_2) &= (\sigma_1 \sigma_3 \sigma_4) + i(\sigma_1 \sigma_3)(\sigma_2 \sigma_4) - i(\sigma_1 \sigma_4)(\sigma_2 \sigma_3) \\ (\sigma_1 \sigma_2 \sigma_3)^2 &= 6 - 2(\sigma_1 \sigma_2) - 2(\sigma_1 \sigma_3) - 2(\sigma_2 \sigma_3) \\ (\sigma_1 \sigma_2 \sigma_3)(\sigma_1 \sigma_2 \sigma_4) &= -(\sigma_1 \sigma_3)(\sigma_2 \sigma_4) - (\sigma_1 \sigma_4)(\sigma_2 \sigma_3) + 2(\sigma_3 \sigma_4) + i(\sigma_1 \sigma_3 \sigma_4) + i(\sigma_2 \sigma_3 \sigma_4) \\ (\sigma_1 \sigma_2 \sigma_3)(\sigma_1 \sigma_4 \sigma_5) &= +(\sigma_2 \sigma_4)(\sigma_3 \sigma_5) - (\sigma_2 \sigma_5)(\sigma_3 \sigma_4) - i(\sigma_1 \sigma_2)(\sigma_3 \sigma_4 \sigma_5) + i(\sigma_1 \sigma_3)(\sigma_2 \sigma_4 \sigma_5), \end{aligned} \quad (14)$$

where $(\sigma_1 \sigma_2 \sigma_3)$ is the mixed product of sigma matrices. Further, a product of two mixed products can always be represented through scalar products:

$$(\sigma_1 \sigma_2 \sigma_3)(\sigma_4 \sigma_5 \sigma_6) = \det \begin{pmatrix} (\sigma_1 \sigma_4) & (\sigma_2 \sigma_4) & (\sigma_3 \sigma_4) \\ (\sigma_1 \sigma_5) & (\sigma_2 \sigma_5) & (\sigma_3 \sigma_5) \\ (\sigma_1 \sigma_6) & (\sigma_2 \sigma_6) & (\sigma_3 \sigma_6) \end{pmatrix}. \quad (15)$$

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

$$(X, Y) \equiv \text{tr}(X^\dagger Y) \quad (16)$$

(not to be confused with the scalar product of sigma matrices). We consider the case when $X, Y \in \{A_{\mathcal{A}}\}$. If supports of X and Y on a lattice do not coincide, then $(X, Y) = 0$. If X and Y

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

have the same support, then $(X, Y) = 2^N 3^C$, where C is the number of cycles arising when the bonds contained in X are superimposed on the bonds contained in Y (*cf.* ref. [8]). In particular,

$$\text{tr}((\sigma_i \sigma_j)(\sigma_j \sigma_k) \dots (\sigma_l \sigma_m)(\sigma_m \sigma_i)) = 3 \cdot 2^N \quad (17)$$

(one cycle) and

$$\text{tr}((\sigma_i \sigma_j \sigma_k)(\sigma_i \sigma_j \sigma_n)(\sigma_k \sigma_l) \dots (\sigma_m \sigma_n)) = \text{tr}((\sigma_i \sigma_j \sigma_k)(\sigma_i \sigma_j \sigma_k)) = 6 \cdot 2^N. \quad (18)$$

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

$$\{B_{\mathcal{B}}\} = \{(\sigma_p \sigma_r \sigma_s), (\sigma_p \sigma_r \sigma_s)(\sigma_j \sigma_k), (\sigma_p \sigma_r \sigma_s)(\sigma_j \sigma_k)(\sigma_l \sigma_m), \dots\}$$

Elements of sets $\{A_{\mathcal{A}}\}$ and $\{A_{\mathcal{A}}, B_{\mathcal{B}}\} \equiv A_{\mathcal{A}} \cup B_{\mathcal{B}}$ are not mutually orthogonal. For example,

$$\begin{aligned} X &= (\sigma_1 \sigma_2)(\sigma_3 \sigma_4) \\ Y &= (\sigma_1 \sigma_3)(\sigma_2 \sigma_4) \\ Z &= (\sigma_1 \sigma_4)(\sigma_2 \sigma_3) \\ g &= \begin{pmatrix} (XX) & (XY) & (XZ) \\ (YX) & (YY) & (YZ) \\ (ZX) & (ZY) & (ZZ) \end{pmatrix} = \begin{pmatrix} 9 & 3 & 3 \\ 3 & 9 & 3 \\ 3 & 3 & 9 \end{pmatrix} \end{aligned} \quad (19)$$

Further, the set $\{A_{\mathcal{A}}\}$ (and, consequently, $\{A_{\mathcal{A}}, B_{\mathcal{B}}\}$) is overcomplete. However, we put forward a hypothesis that all linear dependencies within $\{A_{\mathcal{A}}, B_{\mathcal{B}}\}$ can be described by

$$+(\sigma_1 \sigma_2)(\sigma_3 \sigma_4 \sigma_5) - (\sigma_1 \sigma_3)(\sigma_2 \sigma_4 \sigma_5) + (\sigma_1 \sigma_4)(\sigma_2 \sigma_3 \sigma_5) - (\sigma_1 \sigma_5)(\sigma_2 \sigma_3 \sigma_4) = 0 \quad (20)$$

for elements with an odd number of spins, and by

$$\det \begin{pmatrix} (\sigma_1 \sigma_5) & (\sigma_2 \sigma_5) & (\sigma_3 \sigma_5) & (\sigma_4 \sigma_5) \\ (\sigma_1 \sigma_6) & (\sigma_2 \sigma_6) & (\sigma_3 \sigma_6) & (\sigma_4 \sigma_6) \\ (\sigma_1 \sigma_7) & (\sigma_2 \sigma_7) & (\sigma_3 \sigma_7) & (\sigma_4 \sigma_7) \\ (\sigma_1 \sigma_8) & (\sigma_2 \sigma_8) & (\sigma_3 \sigma_8) & (\sigma_4 \sigma_8) \end{pmatrix} = 0 \quad (21)$$

Table 2. Anderson bound compared to the bound (8) for a translation-invariant linear chain with the Heisenberg Hamiltonian (23). Given are the lower bounds on the ground state energy per spin, E_{gs}/N . The exact Bethe ansatz result reads $E_{gs}/N = 1 - 4 \log 2 \simeq -1.77259$.

cluster size	Anderson bound	bound (8)
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

for elements with an even number of spins. Observe that the latter formula is a consequence of eqs. (15) and (20). We tested this hypothesis for up to 10 spins.

The number of elements in the basis $\{A_A\}$ without taking into account overcompleteness is equal to

$$K(N) = \sum_{k=0}^{[N/2]} C_N^{2k} (2k-1)!! , \quad (22)$$

where $[N/2]$ is the integer part of $N/2$. $K(N)$ grows faster than exponentially with N (see table 1), however it can be rather small for $N \sim 10$.

3.3. An example

Here we consider a one-dimensional lattice of N spins $1/2$ with the nearest-neighbour Heisenberg Hamiltonian

$$H = \sum_{j=1}^N (\boldsymbol{\sigma}_j \boldsymbol{\sigma}_{j+1}) , \quad (23)$$

where $\boldsymbol{\sigma}_{N+1} \equiv \boldsymbol{\sigma}_1$ which ensures translation invariance. This system is integrable by means of Bethe ansatz and the ground state energy per spin in the thermodynamic limit reads $E_{gs}/N = 1 - 4 \log 2$ [9]. Our goal is to lower bound E_{gs}/N by means of our method and to compare this bound to the exact value and to the Anderson bound [1].

We start from considering a cluster with 4 spins. Its Hamiltonian reads

$$H_{cl} = (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2) + (\boldsymbol{\sigma}_2 \boldsymbol{\sigma}_3) + (\boldsymbol{\sigma}_3 \boldsymbol{\sigma}_4) \quad (24)$$

The basis supported by the cluster reads

$$\begin{aligned} \{A_k\} = \{ & 1, (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2), (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_3), (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_4), (\boldsymbol{\sigma}_2 \boldsymbol{\sigma}_3), (\boldsymbol{\sigma}_2 \boldsymbol{\sigma}_4), (\boldsymbol{\sigma}_3 \boldsymbol{\sigma}_4), \\ & (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2)(\boldsymbol{\sigma}_3 \boldsymbol{\sigma}_4), (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_3)(\boldsymbol{\sigma}_2 \boldsymbol{\sigma}_4), (\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_4)(\boldsymbol{\sigma}_2 \boldsymbol{\sigma}_3) \}, \quad k = 0, 1, \dots, 9. \end{aligned} \quad (25)$$

We apply the squaring parametrization [7] in the form

$$\rho_{cl} = \tau^2. \quad (26)$$

Here the reduced density matrix, ρ_{cl} , and the axillary matrix, τ , are expanded in the basis (25) as

$$\rho_{cl} = 2^{-4} a_k A_k, \quad \tau = b_k A_k, \quad (27)$$

where summation over repeating indices is implied. The normalization of ρ_{cl} is imposed by the constraint

$$a_0 = 1. \quad (28)$$

The squaring parametrization (26) implies that each coefficient a_k is a quadratic function of coefficients b_k .

Translational invariance implies

$$a_{(1,2)} = a_{(2,3)} = a_{(3,4)}, \quad a_{(1,3)} = a_{(2,4)}. \quad (29)$$

Here we use self-explanatory notations for indexes, e.g. $a_{(1,2)}$ is the coefficient in front of $(\sigma_1 \sigma_2)$. The Hamiltonian (24) 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:

$$b_{(1,2)} = b_{(3,4)}, \quad b_{(1,3)} = b_{(2,4)} \quad \Rightarrow \quad a_{(1,2)} = a_{(3,4)}, \quad a_{(1,3)} = a_{(2,4)}. \quad (30)$$

The condition

$$a_{(1,2)} = a_{(2,3)} \quad (31)$$

remains and should be accounted for during optimization.

Finally we perform a numerical search for a minimum of $\text{tr } H_{cl} \tau^2$ with the constraints (28) and (31), which are equivalent to $\text{tr } \tau^2 = 1$ and $\text{tr}(\sigma_1, \sigma_2) \tau^2 = (\sigma_2, \sigma_3) \tau^2$, respectively. The resulting bound is presented in table 2, along with the analogous bounds with other cluster sizes. One can see that for a given cluster size the bound (8) outperforms the Anderson bound. The caveat here is that for a given cluster size the calculations for our bound (8) require much more resources than those for the Anderson bound. Whether the bound (8) is able to compete with the Anderson bound in practical numerical computations is a question open for future research.

4. Summary

We have derived a variational *lower* bound (8) on the ground state energy of a quantum system possessing symmetries. The variation should be performed over a certain set of reduced density matrices. Technically, the variation can be performed by means of the squaring parametrization of density matrices [7] which automatically satisfies the positivity condition. We have discussed how this bound can be applied for translation-invariant spin systems with the Heisenberg interaction. Promising results has been obtained in a simple example of a linear chain, see table 2.

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 *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 Schrödinger 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
- [9] Bethe H 1931 Zur theorie der metalle *Zeitschrift für Physik* **71** 205