Polynomial-time algorithm for simulation of weakly interacting quantum spin systems

Sergey Bravyi[(1)]*, David DiVincenzo[(1)], and Daniel Loss[(2)]

[(1)] *IBM T.J. Watson Research Center, Yorktown Heights NY 10598, USA.*
[(2)] *Department of Physics, Klingelbergstrasse 82, University of Basel, 4056 Basel, Switzerland*

March 5, 2022

Abstract

We describe an algorithm that computes the ground state energy and correlation functions for 2-local Hamiltonians in which interactions between qubits are weak compared to single-qubit terms. The running time of the algorithm is polynomial in n and δ^{-1}, where n is the number of qubits, and δ is the required precision. Specifically, we consider Hamiltonians of the form $H = H_0 + \epsilon V$, where H_0 describes non-interacting qubits, V is a perturbation that involves arbitrary two-qubit interactions on a graph of bounded degree, and ϵ is a small parameter. The algorithm works if $|\epsilon|$ is below a certain threshold value ϵ_0 that depends only upon the spectral gap of H_0, the maximal degree of the graph, and the maximal norm of the two-qubit interactions. The main technical ingredient of the algorithm is a generalized Kirkwood-Thomas ansatz for the ground state. The parameters of the ansatz are computed using perturbative expansions in powers of ϵ. Our algorithm is closely related to the coupled cluster method used in quantum chemistry.

---

*e-mail: sbravyi@us.ibm.com

Contents

1   Introduction and summary of results

Perturbation theory provides a systematic way of getting approximations to eigenvalues and eigenvectors for a variety of quantum spin models. Arguably, a significant part of analytical and numerical results of condensed matter physics has been obtained using perturbative expansions in some small parameter. Quite recently the methods of the perturbation theory have been successfully employed in quantum complexity theory. In Ref. [1] Kempe, Kitaev, and Regev used perturbative reductions to show that the problem of computing the ground state energy of a Hamiltonian with two-qubit interactions is QMA-complete. After that Terhal and Oliveira [2] generalized this result to local Hamiltonians on a 2D lattice.

Our main goal is to examine whether the methods of the perturbation theory provide an efficient computational algorithm for the simulation of quantum spin systems. In this paper we focus on the simulation of low-temperature properties, namely computing the ground state energy and spin-spin correlation functions for the ground state. An efficient algorithm must have a running time $T = O(n^\alpha \delta^{-\beta})$, where n is the number of spins, δ is a precision up to which we need to compute the ground state energy or a correlation function, and $\alpha, \beta > 0$ are some constants.

Before stating the results, let us describe the spin models that we shall consider. Let $\mathcal{G} = (\mathcal{L}, \mathcal{E})$ be a graph with a set of vertices $\mathcal{L}$, $|\mathcal{L}| = n$, and set of edges $\mathcal{E}$. Suppose n spins-1/2 (qubits) are located at vertices $u \in \mathcal{L}$ and spin-spin interactions are located on edges $(u, v) \in \mathcal{E}$. The Hamiltonian is

$$H(\epsilon) = H_0 + \epsilon V, \qquad H_0 = \sum_{u \in \mathcal{L}} \Delta_u \, |1\rangle\langle 1|_u, \qquad V = \sum_{(u,v) \in \mathcal{E}} V_{u,v}. \tag{1}$$

Here $V_{u,v}$ is an arbitrary operator acting on a pair of qubits u, v, and ϵ is a real number. The operators H_0 and V are called the *unperturbed Hamiltonian* and the *perturbation*. We shall always assume that $\Delta_u > 0$ for all $u \in \mathcal{L}$. Accordingly, the unperturbed Hamiltonian H_0 has a non-degenerate ground state

$$|\Omega\rangle = |0, 0, \ldots, 0\rangle, \quad H_0 \, |\Omega\rangle = 0.$$

Most of the time, all we will need to know about H_0 and V are the following parameters

$$\Delta = \min_{u \in \mathcal{L}} \Delta_u, \quad J = \max_{(u,v) \in \mathcal{E}} \|V_{u,v}\|. \tag{2}$$

The parameter Δ is the gap between the smallest and the second smallest eigenvalue of H_0, while the parameter J characterizes a strength of the perturbation V. Let d be the maximum vertex degree of the graph $\mathcal{G}$,

$$d = \max_{u \in \mathcal{L}} |\{v \, : \, (u, v) \in \mathcal{E}\}| \, . \tag{3}$$

The quantity we are interested in is the smallest eigenvalue of $H(\epsilon)$, which we shall denote by $E(\epsilon)$. Clearly, $E(\epsilon)$ is a continuous concave function of ϵ and $E(0) = 0$. Besides, since we assume that $\Delta > 0$, the standard perturbation theory arguments [3] show that $E(\epsilon)$ is analytic at $\epsilon = 0$ and the Taylor series

$$E(\epsilon) = \sum_{p=1}^{\infty} E_p \, \epsilon^p \tag{4}$$

converges absolutely for $\|\epsilon V\| < \Delta/2$. The following theorem proved by Yarotsky [4] asserts that $E(\epsilon)$ is non-degenerate for sufficiently small ϵ and sets a lower bound on the spectral gap.

Theorem 1. *Suppose $|\epsilon| \le 2\epsilon_0$, where*

$$\epsilon_0 = \frac{2^{-18}\,\Delta}{dJ}. \tag{5}$$

Then the smallest eigenvalue $E(\epsilon)$ has multiplicity 1 and the gap between $E(\epsilon)$ and the second smallest eigenvalue of $H(\epsilon)$ is at least $\Delta/2$.

(The explicit value of ϵ_0 has not been stated in Ref. [4].) We shall provide an alternative proof of Theorem 1 in Sections 2,3.

As was shown by Osborne in Ref. [5], Theorem 1 implies that expectation values of local observables on the ground state of $H(\epsilon)$ can be efficiently computed within any constant precision δ by simulating quantum adiabatic evolution along the path connecting $H(0)$ and $H(\epsilon)$. However, the running time of such simulation scales exponentially as a function of δ^{-1}. As was noted in Ref. [5], it means that simulation of the adiabatic evolution does not yield a polynomial-time algorithm for computing the ground state energy.

The perturbation theory provides an approximation to the ground state energy by truncating the series Eq. (4) at sufficiently high order p. In order to understand whether this approach can be used to construct an efficient computational algorithm, two separate issues have to be addressed:

Q1: What is the convergence radius of the perturbative series?

Q2: What is the computational cost of finding the coefficients in the perturbative series?

Note that the radius of convergence of the series Eq. (4) is a property of the Hamiltonian $H(\epsilon)$ only. It does not depend upon what particular perturbative expansion has been used to find the coefficients E_p. The following theorem allows one to answer the first question.

Theorem 2. *The Taylor series $E(\epsilon) = \sum_{p=1}^{\infty} E_p\,\epsilon^p$ converges absolutely for $|\epsilon| \le 2\epsilon_0$. Furthermore,*

$$\left| E(\epsilon) - \sum_{q=1}^{p} E_q\,\epsilon^q \right| \le n\Delta 2^{-16-p} \quad if \quad |\epsilon| \le \epsilon_0. \tag{6}$$

Thus if one needs to compute $E(\epsilon)$ with a specified precision δ, it suffices to compute the coefficients $E_1, \ldots, E_p$, where $p = \log_2(n\delta^{-1}) + O(1)$ (assuming that Δ is a constant that does not depend on n).

Answering the second question has nothing to do with the convergence radius of the series Eq. (4) (as long as it is non-zero). One can compute the coefficients E_p by choosing ϵ so small that $\|\epsilon V\| \ll \Delta$. In this regime the standard perturbation theory is applicable, for example, the self-energy operator formalism, see Refs. [1, 6], or the Rayleigh-Schrödinger expansion, see Ref. [7]. Clearly, the computational cost of finding the coefficients E_p varies for different methods.

In the present paper we compute the coefficients E_p using the Kirkwood-Thomas ansatz for the ground state. It was originally proposed in Ref. [8] for translation-invariant Ising-like Hamiltonians with a transverse magnetic field. The translation-invariance constraint has been removed in the later work by Datta and Kennedy [9]. We use the generalized Kirkwood-Thomas ansatz proposed by Yarotsky [4] which is applicable to any spin Hamiltonian with sufficiently weak interactions. It allows us to prove the following.

Theorem 3. *Suppose d is a fixed constant independent of n. Then there exists an algorithm with a running time $n \exp(O(p))$ that takes as input a triple (H_0, V, p) and outputs $E_1, \ldots, E_p$.*

An immediate consequence of Theorems 2,3 is

Corollary 1. *Suppose Δ, J, d are fixed constants independent of n and $|\epsilon| \le \epsilon_0$. Then there exists an algorithm with a running time $poly(n, \delta)$ that computes $E(\epsilon)$ with an absolute error at most δ.*

Besides, it follows from Theorems 2,3 that the energy *density* $E(\epsilon)/n$ can be computed with a precision δ in a time $n \cdot poly(\delta^{-1})$.

Note that while computing the coefficients $E_1, \ldots, E_p$ we cannot afford the running time to grow faster than $\exp(O(p))$ (for fixed n) since we need $p \sim \log(n\delta^{-1})$ to achieve the desired accuracy. The perturbative expansion based on the Kirkwood-Thomas ansatz has two special features that make the scaling $\exp(O(p))$ possible: (i) The parameters of the ansatz are complex amplitudes $C(M)$ assigned to subsets of vertices $M \subseteq \mathcal{L}$. The recursive equations specifying the amplitudes $C(M)$ are described by a polynomial of a *constant degree*, see Section 3.1; (ii) The perturbative expansion $C(M) = \sum_{p=1}^{\infty} C_p(M) \epsilon^p$ has a property known as the linked cluster theorem, namely, $C_p(M) = 0$ unless M can be covered by a connected subgraph of size $O(p)$, see Section 4.1. The number of such subgraphs grows only exponentially with p, see Section 4.2. It implies that the number of non-zero coefficients $C_p(M)$ grows as $n \exp(O(p))$, see Section 5.1.

Naturally, one could run the algorithm from Theorem 3 to compute the truncated series for $E(\epsilon)$ even if $|\epsilon| > \epsilon_0$. The running time will be polynomial in n and δ^{-1} as long as $|\epsilon|$ is smaller than the convergence radius R of the series Eq. (4). Although we believe that R must be close to $\Delta/(dJ)$ (see a discussion at Section 6), its exact value cannot be easily found. In practical simulations, one could evaluate R by computing sufficiently many coefficients E_p and using the fact that R^{-1} is the largest accumulation point of a sequence $|E_p|^{1/p}$, $p = 1, \ldots, \infty$, see Ref. [10]. Note that in general the singular point (points) of $E(\epsilon)$ with $|\epsilon| = R$ does not lie on the real axis and thus cannot be identified with a quantum phase transition point of $H(\epsilon)$ (since we consider finite systems, the latter is not even well defined).

Obviously, efficient computation of $E(\epsilon)$ is possible due to the presence of a small parameter ϵ in the problem. However it should be emphasized that the condition $|\epsilon| \le \epsilon_0$ does not imply that the ground state $|\psi\rangle$ of $H(\epsilon)$ is close to the ground state $|\Omega\rangle$ of the unperturbed Hamiltonian H_0. In fact, one should expect that $|\psi\rangle$ and $|\Omega\rangle$ are almost orthogonal for large n[1]. To illustrate this statement, consider as an example the perturbation $V = -J \sum_{u \in \mathcal{L}} X_u$, where X is the Pauli σ^x operator, and the unperturbed Hamiltonian $H_0 = \Delta \sum_{u \in \mathcal{L}} |1\rangle\langle 1|_u$. Clearly, the ground state of $H(\epsilon)$ is a product of one-qubit states, $|\psi\rangle = \bigotimes_{u \in \mathcal{L}} |\psi_u\rangle$. A simple calculation shows that $\langle 0|\psi_u\rangle = \cos(\theta/2)$, where $\cos(\theta) = (1 + 4\epsilon^2 J^2/\Delta^2)^{-1/2}$. Thus for any fixed ϵ the overlap $\langle\Omega|\psi\rangle = (\cos(\theta/2))^n$ gets exponentially small as n increases. However the reduced density matrices of the ground states $|\psi\rangle$ and $|\Omega\rangle$ for any subset of qubits of constant size are indeed close to each other for small ϵ. In other words, for small ϵ the state $|\psi\rangle$ describes small density quantum fluctuations of the background state $|\Omega\rangle$. One could speculate that this statement remains true for arbitrary weak perturbations as well. The Kirkwood-Thomas ansatz for the ground state of $H(\epsilon)$ used in the present paper provides a convenient way to quantify the "density of quantum fluctuations" and prove that it is indeed small for $|\epsilon| \le \epsilon_0$. Unfortunately, our approach does not allow us to make any statements about the validity of the area

---

[1]This effect is analogous to the well-known "orthogonality catastrophe" observed by Anderson in Ref. [11] for non-interacting fermions in a presence of a scattering potential.

law or to decide whether the ground state can be well approximated using the PEPS ansatz, see Ref. [12].

The Kirkwood-Thomas ansatz is not well suited for computing spin-spin correlation functions because it provides an unnormalized ground state. We avoid this problem using the standard relation between the correlation functions and the linear response of the ground state energy to a small perturbation. It allows us to prove

Theorem 4. *Let $O_{u,v}$ be a Hermitian operator acting non-trivially only on qubits $u, v \in \mathcal{L}$. Suppose $\|O_{u,v}\| \leq 1$. The expectation value of $O_{u,v}$ on the ground state of $H(\epsilon)$ can be computed with a precision δ in a time $T = poly(\delta^{-1})$ as long as $|\epsilon| \leq \epsilon_0/2(d+1)$.*

Remark: After the present work has been completed, it was communicated to us by F. Verstraete [13] that the simulation algorithm based on the Kirkwood-Thomas ansatz is closely related to the coupled cluster method originally introduced by Coester [14]. The coupled cluster method is extensively used for numerical simulations in quantum chemistry, see a review [15], as well as in condensed matter physics, see a review [16] and the references therein. Accordingly, from the perspective of practical simulations, the algorithm described in the present paper is certainly not a new one. However, we believe that our results provide the first rigorous proof that the coupled cluster method yields a polynomial-time simulation algorithm for spin Hamiltonians with weak interactions.

The rest of the paper is organized as follows. Section 2 provides the necessary background on the generalized Kirkwood-Thomas ansatz. It mostly follows Ref. [4], although some of our proofs are technically different (in particular, Lemma 4). Section 3 shows how to solve the Kirkwood-Thomas equations using a power series and proves Theorem 2. In Section 4 we prove that our perturbative expansion obeys the well-known linked cluster theorem and establish an upper bound on the number of linked clusters on a graph. The algorithms for computing the ground state energy and spin-spin correlation functions are explicitly described in Section 5 which provides a proof of Theorems 3,4. Some open problems are discussed in Section 6. Appendix A contains a technical lemma proving submultiplicativity of the norm of creation operators.

2 Kirkwood-Thomas ansatz for the ground state

2.1 Creation operators

Define one-qubit operator $a^\dagger = |1\rangle\langle 0|$. Let $a_u^\dagger$ be the operator $a^\dagger$ on qubit u tensored with the identity on all other qubits. For any non-empty subset of vertices $M \subseteq \mathcal{L}$ denote $a_M^\dagger = \prod_{u \in M} a_u^\dagger$. Note that the operators $a_M^\dagger$ are nilpotent, $(a_M^\dagger)^2 = 0$, and that they pairwise commute: $a_M^\dagger a_K^\dagger = a_K^\dagger a_M^\dagger$. Also, one can easily check that the operators $\{a_M^\dagger\}$, $\emptyset \neq M \subseteq \mathcal{L}$ are linearly independent. (All these definitions and properties apply to a and a_M operators as well).

Definition 1. *A creation operator is an operator that can be written as*

$$C = \sum_{\emptyset \neq M \subseteq \mathcal{L}} C(M) \, a_M^\dagger$$

for some complex numbers $C(M)$.

For any given creation operator C the coefficients $C(M)$ are uniquely defined by $C(M) = \langle \Omega | a_M C | \Omega \rangle$.

Claim 1. *Any state $|\psi\rangle$ satisfying $\langle \Omega | \psi \rangle = 1$ can be uniquely written as $|\psi\rangle = exp(-C) |\Omega\rangle$ for some creation operator C.*

Remark: the exponent above is defined by its Taylor series. The nilpotence of operators $a_M^\dagger$ implies that $C^k = 0$ for any k greater than the number of qubits $n = |\mathcal{L}|$, so the Taylor series can be truncated at $k = n$.

Proof: Clearly, the states $\{a_M^\dagger |\Omega\rangle\}$, $M \subseteq \mathcal{L}$, constitute the orthonormal basis of the n-qubit Hilbert space. Let $|\psi\rangle = \sum_{M \subseteq \mathcal{L}} \psi(M) a_M^\dagger |\Omega\rangle$. Equation $|\psi\rangle = \exp(-C) |\Omega\rangle$ is equivalent to a system of equations

$$\psi(\emptyset) = 1, \quad C(M) = -\psi(M) + \sum_{k=2}^{|M|} \frac{(-1)^k}{k!} \sum_{M = M_1 \cup \ldots \cup M_k} C(M_1) \cdots C(M_k), \quad M \subseteq \mathcal{L}, \quad M \neq \emptyset. \quad (7)$$

Here the second summation is over all partitions of M into k disjoint non-empty sets $M_1, \ldots, M_k$. Suppose we have already found all coefficients $C(M)$ with $|M| \leq p$. Then Eq. (7) assigns a unique value to all coefficients $C(M)$ with $|M| = p + 1$. Thus the system Eq. (7) has a unique solution. $\square$

2.2 Ansatz for the ground state

Our goal is to find an eigenvector $|\psi\rangle$ satisfying $H(\epsilon) |\psi\rangle = E(\epsilon) |\psi\rangle$, where $E(\epsilon)$ is the smallest eigenvalue of $H(\epsilon)$. We shall use the following ansatz for $|\psi\rangle$ (we don't care about the normalization):

$$|\psi\rangle = \exp(-C) |\Omega\rangle, \quad C = \sum_{\emptyset \neq M \subseteq \mathcal{L}} C(M) a_M^\dagger. \quad (8)$$

Claim 1 asserts that the ground state can be represented in this form unless it is orthogonal to $|\Omega\rangle$. Since we don't require $|\psi\rangle$ to be a normalized state, the ansatz Eq. (8) is meaningful only if C is a bounded operator. We shall define a norm of a creation operator as

$$\|C\|_1 = \max_{u \in \mathcal{L}} \sum_{M \ni u} |C(M)|. \quad (9)$$

Thus the ansatz Eq. (8) must be supplemented by a requirement that C is a creation operator with a finite norm $\|C\|_1$. Of course, it may happen that H has several eigenvectors of the form Eq. (8). One has to invoke some extra arguments to select an eigenvector corresponding to the smallest eigenvalue, see subsection 2.4.

The "physical meaning" of the norm $\|C\|_1$ can be illustrated by considering a product state: $|\psi\rangle = \bigotimes_{u \in \mathcal{L}} |\psi_u\rangle$, where $|\psi_u\rangle = |0\rangle + \alpha_u |1\rangle$. Obviously, $|\psi\rangle = \exp(-C) |\Omega\rangle$ with $C = -\sum_{u \in \mathcal{L}} \alpha_u a_u^\dagger$. Accordingly, $\|C\|_1 = \max_{u \in \mathcal{L}} |\alpha_u|$. Thus one can think about $\|C\|_1$ as a density of quantum fluctuations.

Using the identity $\exp(C) \exp(-C) = I$ valid for arbitrary operator C, see [20], one can rewrite the Schrödinger equation $H(\epsilon) |\psi\rangle = E(\epsilon) |\psi\rangle$ as

$$\exp(\hat{C})(H_0)|\Omega\rangle + \epsilon \exp(\hat{C})(V)|\Omega\rangle = E(\epsilon) |\Omega\rangle. \quad (10)$$

Here we introduced a superoperator[2] $\hat{C}$ such that

$$\hat{C}(X) = CX - XC.$$

The exponent $\exp{(\hat{C})}$ is defined by the Taylor series. The advantage of the ansatz Eq. (8) is that we can truncate expansion of the exponent $\exp{(\hat{C})}$ after a few lowest orders since all higher order terms turn out to be identically zero. It follows from the two lemmas stated below.

Lemma 1. *Let C_1, C_2 be creation operators. Then*

$$\hat{C}_1 \hat{C}_2(H_0) = 0. \tag{11}$$

Proof: To simplify notations we shall consider operators a instead of $a^\dagger$. Let $u, M_1, M_2 \subseteq \mathcal{L}$ and $X = \hat{a}_{M_1} \hat{a}_{M_2}(|1\rangle\langle 1|_u)$. By linearity, it is enough to prove that $X = 0$. Since the operators a_{M_1} and a_{M_2} commute, $X = 0$ unless $u \in M_1 \cap M_2$. Then $[a_{M_2}, |1\rangle\langle 1|_u] = a_{M_2}$ and $X = [a_{M_1}, a_{M_2}] = 0$.

$\square$

Lemma 2. *Let $C_1, C_2, \ldots, C_5$ be creation operators. Then*

$$\hat{C}_1 \hat{C}_2 \cdots \hat{C}_5(V) = 0. \tag{12}$$

Proof: To simplify notations we shall consider operators a instead of $a^\dagger$. Let $M_1, M_2, \ldots, M_5 \subseteq \mathcal{L}$, $(u, v) \in \mathcal{E}$ and $X = \hat{a}_{M_1} \hat{a}_{M_2} \cdots \hat{a}_{M_5}(V_{u,v})$. By linearity it is enough to prove that $X = 0$. Since the operators $a_{M_1}, \ldots, a_{M_5}$ commute with each other, $X = 0$ unless each of the subsets M_j contains at least one of the vertices u, v. Therefore, expanding the commutators one can represent X as a linear combination of 2^5 terms, where each term contains at least five operators a_u, a_v on the pair of qubits u, v. Some of these operators a are on the right of $V_{u,v}$ and some of them are on the left. Thus at least three operators a are on the same side of $V_{u,v}$. Then at least two operators a act on the same side of $V_{u,v}$ and on the same qubit. Thus each of the 2^5 terms in X contains either a_u^2 or a_v^2. Thus $X = 0$.

$\square$

Combining Lemmas 1,2 we get the following truncations:

$$\exp{(\hat{C})}(H_0) = H_0 + \hat{C}(H_0), \tag{13}$$

$$\exp{(\hat{C})}(V) = \sum_{k=0}^{4} \frac{1}{k!} \hat{C}^k(V). \tag{14}$$

Here a convention $\hat{C}^0(V) = V$ is adopted.

Let us point out an analogy between the truncation effect observed above and the Lieb-Robinson bound [17, 18]. The latter asserts that for any local observable O_u acting only on a qubit u and for any Hamiltonian H with short-range interactions of bounded norm the time evolved observable $O_u(t) = \exp{(i\hat{H}t)}(O_u)$ can be approximated very well by an operator acting only on spins within

---

[2]In our context a superoperator is a linear operator acting on the space of linear operators on $\mathcal{H}$. Throughout the paper we shall use a notation $\hat{A}$ for a superoperator $\hat{A}(X) = AX - XA$ associated with a linear operator A.

distance $v|t|$ from u, where v is a group velocity. If one takes a creation operator C for which the coefficients $C(M)$ are non-zero only for subsets M of size $O(1)$ (an analogue of short-range interactions), then the "time-evolved" observable $\exp(\hat{C})(O_u)$ acts only on the spins within distance $O(1)$ from u (apply the same arguments as in the proof of Lemma 2). As opposed to the Lieb-Robinson bound scenario, the size of a region acted on by the evolved operator does not depend on the norm of C (which is analogous to the evolution time) and no approximations are involved.

2.3 Kirkwood-Thomas equations

Substituting Eqs. (13) into the Schrödinger equation Eq. (10) and taking into account that $H_0 |\Omega\rangle = 0$ one gets

$$- \sum_{\emptyset \neq M \subseteq \mathcal{L}} C(M) H_0\, a_M^\dagger |\Omega\rangle + \epsilon \exp(\hat{C})(V) |\Omega\rangle = E(\epsilon) |\Omega\rangle. \tag{15}$$

Let us introduce eigenvalues of the unpertubed Hamiltonian $E_0(M)$ such that

$$H_0\, a_M^\dagger |\Omega\rangle = E_0(M)\, a_M^\dagger |\Omega\rangle, \quad E_0(M) = \sum_{u \in M} \Delta_u. \tag{16}$$

Multiplying Eq. (15) on the left by $\langle \Omega | a_M$, $M \neq \emptyset$, and employing Eq. (14) one arrives at

$$C(M) = \frac{\epsilon}{E_0(M)} \sum_{k=0}^{4} \frac{1}{k!} \langle \Omega | a_M \hat{C}^k(V) | \Omega \rangle, \quad \emptyset \neq M \subseteq \mathcal{L}. \tag{17}$$

Following [4], we shall refer to Eq. (17) as *Kirkwood-Thomas equations*. Similarly, multiplying Eq. (15) by $\langle \Omega |$ on the left one gets

$$E(\epsilon) = \epsilon \sum_{k=0}^{4} \frac{1}{k!} \langle \Omega | \hat{C}^k(V) | \Omega \rangle. \tag{18}$$

It is clear that the Kirkwood-Thomas equations Eq. (17) may have several solutions C since the equations do not explicitly include the eigenvalue $E(\epsilon)$. In the worst case when neither eigenvector of $H(\epsilon)$ is orthogonal to $|\Omega\rangle$ the Kirkwood-Thomas equations would have 2^n solutions since any eigenvector could be represented in the form Eq. (8). We shall explain how to select the solution corresponding to the smallest eigenvalue in the next subsection.

The following lemma asserts that the norm $\|\cdot\|_1$ has a property analogous to submultiplicativity. It is the main technical tool that allows one to manipulate easily with equations like Eq. (17).

Lemma 3. *Let k be any integer and $C_1, \ldots, C_k$ be creation operators. Define a creation operator C such that*

$$C = \sum_{\emptyset \neq M \subseteq \mathcal{L}} C(M) a_M^\dagger \quad where \quad C(M) = \frac{1}{E_0(M)} \langle \Omega | a_M \hat{C}_1 \cdots \hat{C}_k(V) | \Omega \rangle.$$

Then

$$\|C\|_1 \leq \frac{2^{13} d J}{\Delta} \prod_{j=1}^{k} \|C_j\|_1. \tag{19}$$

9

Besides,

$$|\langle\Omega|\hat{C}_1\cdots\hat{C}_k(V_{u,v})|\Omega\rangle| \leq 2^4 J \prod_{j=1}^{k} \|C_j\|_1 \quad \textit{for any} \quad (u,v) \in \mathcal{E}. \tag{20}$$

The proof of the lemma is presented in Appendix A.

2.4   A lower bound on the spectral gap

Suppose we can find *some* eigenvalue $E'(\epsilon)$ of the Hamiltonian $H(\epsilon)$ such that $E'(0) = 0$, $E'(\epsilon)$ is a continuous function of ϵ, and $E'(\epsilon)$ has multiplicity 1 for $|\epsilon| \leq \epsilon_c$. Then it follows immediately that $E'(\epsilon)$ is the *smallest* eigenvalue of $H(\epsilon)$ for all $|\epsilon| \leq \epsilon_c$. Of course, the main difficulty in using this argument is proving non-degeneracy of an eigenvalue. The following lemma asserts that a solution of the Kirkwood-Thomas equations Eq. (17) with a sufficiently small norm $\|C\|_1$ corresponds to a non-degenerate eigenvalue separated from the rest of the spectrum by a constant gap.

Lemma 4. *Suppose $H(\epsilon)|\psi\rangle = E(\epsilon)|\psi\rangle$, where $|\psi\rangle = \exp(-C)|\Omega\rangle$ and C is a creation operator with a finite norm $\|C\|_1$ satisfying the inequality*

$$1 > \frac{2^{14}dJ|\epsilon|}{\Delta}\sum_{k=0}^{3}\frac{(\|C\|_1)^k}{k!}. \tag{21}$$

Then $E(\epsilon)$ has multiplicity 1 and any other eigenvalue of $H(\epsilon)$ is separated from $E(\epsilon)$ by a gap at least $\Delta/2$.

    Proof: Let us abbreviate $H \equiv H(\epsilon)$. Assume that $H|\phi\rangle = (E(\epsilon) + \delta)|\phi\rangle$ where $|\delta| < \Delta/2$ and the states $|\psi\rangle$, $|\phi\rangle$ are linearly independent (the latter condition is fulfilled automatically if $\delta \neq 0$). We can always write $|\phi\rangle$ as

$$\exp(C)|\phi\rangle = \sum_{M\subseteq\mathcal{L}} B(M)a_M^\dagger|\Omega\rangle \tag{22}$$

for some complex numbers $B(M)$. Note that $B(M) \neq 0$ for some non-empty set M since otherwise $|\phi\rangle$ is proportional to $|\psi\rangle$. Thus we can define a creation operator $B = \sum_{\emptyset\neq M\subseteq\mathcal{L}} B(M)a_M^\dagger$ with a non-zero norm $\|B\|_1 > 0$. Using commutativity $[C,B] = 0$ we can represent $|\phi\rangle$ as

$$|\phi\rangle = B|\psi\rangle + B(\emptyset)|\psi\rangle.$$

Then the eigenvalue equations $H|\phi\rangle = (E(\epsilon) + \delta)|\phi\rangle$ and $H|\psi\rangle = E(\epsilon)|\psi\rangle$ imply

$$[B,H]|\psi\rangle = [B, H - E(\epsilon)I]|\psi\rangle = B(H - E(\epsilon)I)|\psi\rangle - (H - E(\epsilon)I)|\phi\rangle = -\delta|\phi\rangle = -\delta B|\psi\rangle - \delta B(\emptyset)|\psi\rangle. \tag{23}$$

Commutativity $[C,B] = 0$ yields $\exp(\hat{C})(B) = B$. Hence, multiplying Eq. (23) by $\exp(C)$ on the left one arrives at

$$[B, \exp(\hat{C})(H)]|\Omega\rangle + \delta B|\Omega\rangle + \delta B(\emptyset)|\Omega\rangle = 0. \tag{24}$$

From Lemma 1 we know that $[B, \exp(\hat{C})(H_0)] = [B, H_0]$. Choosing any $M \neq \emptyset$ and multiplying Eq. (24) by $\langle\Omega|a_M$ on the left one gets

$$B(M) = \frac{\epsilon}{E_0(M) - \delta}\langle\Omega|a_M\hat{B}\exp(\hat{C})(V)|\Omega\rangle = \frac{\epsilon E_0(M)}{E_0(M) - \delta}\sum_{k=0}^{3}\frac{1}{k!}\frac{1}{E_0(M)}\langle\Omega|a_M\hat{B}\hat{C}^k(V)|\Omega\rangle.$$

10

Here we have taken into account that $\hat{B}\hat{C}^k(V) = 0$ for $k \geq 4$, see Lemma 2. Note that condition $|\delta| < \Delta/2$ implies a bound $|E_0(M)/(E_0(M) - \delta)| \leq 2$. Applying Lemma 3 to the operator B and using the triangle inequality for the norm one gets

$$\|B\|_1 \leq \frac{2^{14}dJ|\epsilon|}{\Delta}\|B\|_1 \sum_{k=0}^{3} \frac{1}{k!}(\|C\|_1)^k.$$

Since $\|B\|_1 > 0$ we can divide both sides by $\|B\|_1$ getting an inequality opposite to the one stated in the lemma. Thus the assumption from which we started the proof leads to a contradiction.

$\square$

Remark: Note that at $\epsilon = 0$ the Hamiltonian $H(\epsilon) = H_0$ has many degenerate eigenvalues, so one can certainly find two eigenvalues with separation $|\delta| < \Delta/2$. It might seem to be in contradiction with the lemma above. However at $\epsilon = 0$ the condition that $\|C\|_1$ is finite can not be fulfilled for degenerate eigenvalues, since the corresponding eigenvectors are orthogonal to $|\Omega\rangle$.

Corollary 2. *Suppose $H(\epsilon)$ has an eigenvector $|\psi\rangle = \exp(-C)|\Omega\rangle$ with an eigenvalue $E'(\epsilon)$ such that $E'(\epsilon)$ is a continuous function of ϵ, $E'(0) = 0$, and $\|C\|_1 \leq c_{max}$ for all $|\epsilon| \leq \epsilon'_c$. Define ϵ''_c such that*

$$1 = \frac{2^{14}dJ\epsilon''_c}{\Delta} \sum_{k=0}^{3} \frac{1}{k!}(c_{max})^k.$$

Let $\epsilon_c = \min(\epsilon'_c, \epsilon''_c)$. Then for all $|\epsilon| \leq \epsilon_c$
(1) $E'(\epsilon)$ is the smallest eigenvalue of $H(\epsilon)$
(2) $E'(\epsilon)$ has multiplicity 1
(3) $E'(\epsilon)$ is separated from the rest of the spectrum by a gap at least $\Delta/2$

Proof: (1) Indeed, Lemma 4 implies that no level crossings involving the eigenvector $|\psi\rangle$ can occur for $|\epsilon| \leq \epsilon_c$. Since $|\psi\rangle$ is the ground state for $\epsilon = 0$, it is the ground state for all $|\epsilon| \leq \epsilon_c$. (2) and (3) follow immediately from Lemma 4.

$\square$

3 Solution of the Kirkwood-Thomas equations

In their original paper [8] Kirkwood and Thomas employed the expansion in powers of ϵ in order to find the ground state. Alternative approach proposed by Datta and Kennedy in Ref. [9] and generalized by Yarotsky in Ref. [4] is to regard equation Eq. (17) as a fixed point equation for a non-linear map on the space of creation operators. One can prove that this map is a contraction in the unit ball (for a properly defined metric) if ϵ is below certain threshold value. Then one can invoke Brouwer fixed point theorem to argue that the unit ball contains a unique fixed point. Although the latter method is more elegant, we adopt the original Kirkwood-Thomas approach based on power series, because it naturally lends itself for getting approximation to the ground state energy with a controllable error.

3.1 Solution by formal power series

Let us first solve Kirkwood-Thomas equation Eq. (17) in terms of formal power series ignoring the convergence issue. Recall that $C = \sum_{M \subseteq \mathcal{L}} C(M) \, a_M^\dagger$, where the sum is over all non-empty sets. Define a series

$$C(M) = \sum_{p=1}^\infty C_p(M) \, \epsilon^p, \quad \emptyset \neq M \subseteq \mathcal{L}. \tag{25}$$

Let us agree that $C_0(M) = 0$ for any M. Define also

$$C_p = \sum_{\emptyset \neq M \subseteq \mathcal{L}} C_p(M) \, a_M^\dagger, \quad \hat{C}_p = \sum_{\emptyset \neq M \subseteq \mathcal{L}} C_p(M) \, \hat{a}_M^\dagger, \tag{26}$$

so that $C = \sum_{p=1}^\infty C_p \, \epsilon^p$ and $\hat{C} = \sum_{p=1}^\infty \hat{C}_p \, \epsilon^p$. Substituting the series Eqs. (25,26) into the Kirkwood-Thomas equation Eq. (17) and equating the coefficients for each power of ϵ one gets

$$C_1(M) = E_0(M)^{-1} \, \langle \Omega | a_M \, V | \Omega \rangle, \tag{27}$$

$$C_p(M) = E_0(M)^{-1} \sum_{k=1}^4 \frac{1}{k!} \sum_{p_1+\ldots+p_k=p-1} \langle \Omega | a_M \, \hat{C}_{p_1} \cdots \hat{C}_{p_k}(V) | \Omega \rangle, \quad p \geq 2. \tag{28}$$

Clearly, the equations above have a unique solution. Substituting Eqs. (25,26) into the formula for the ground state energy Eq. (18) one gets

$$E(\epsilon) = \sum_{p=1}^\infty E_p \, \epsilon^p, \quad E_1 = \langle \Omega | V | \Omega \rangle, \quad E_p = \sum_{k=1}^4 \frac{1}{k!} \sum_{p_1+\ldots+p_k=p-1} \langle \Omega | \hat{C}_{p_1} \cdots \hat{C}_{p_k}(V) | \Omega \rangle, \quad p \geq 2. \tag{29}$$

Of course, formal power series do not represent an actual solution of the Kirkwood-Thomas equations unless we prove their convergence.

3.2 Convergence of C-series

We would like to prove that the series $C = \sum_{p=1}^\infty C_p \, \epsilon^p$ are convergent with respect to the norm Eq. (9) with a non-zero convergence radius. We shall need to get a lower bound on the convergence radius in terms of Δ, d, and J. Clearly, it is enough to analyze convergence of the series

$$\chi(\epsilon) = \sum_{p=1}^\infty \chi_p \, \epsilon^p, \quad \chi_p = \|C_p\|_1. \tag{30}$$

Note that $\|C\|_1 \leq \chi(|\epsilon|)$.

Lemma 5. *The series $\chi(\epsilon) = \sum_{p=1}^\infty \chi_p \, \epsilon^p$ converges absolutely for*

$$|\epsilon| \leq 2\epsilon_0 = \frac{2^{-17}\Delta}{dJ}. \tag{31}$$

Besides, for any ϵ as above one has the following bounds:

$$|\chi(\epsilon)| \leq 2^{-15}, \quad \chi_p \leq \frac{2^{-15}}{(2\epsilon_0)^p} \quad for \quad p \geq 1. \tag{32}$$

Proof: Let us first get an upper bound on $\|C_1\|_1$. From Eq. (27) it clear that $C_1(M) = 0$ unless $M \subseteq \{u, v\}$ for some edge $(u, v) \in \mathcal{E}$. Let $u \in \mathcal{L}$ be the vertex that achieves the maximum in $\|C_1\|_1 = \max_{u \in \mathcal{L}} \sum_{M \ni u} |C_1(M)|$. Then the sum contains at most $d + 1$ sets M, namely, $M = \{u\}$ and $M = \{u, v\}$ for $(u, v) \in \mathcal{E}$. Therefore $\|C_1\|_1 \leq (d+1)J/\Delta \leq 2dJ/\Delta$, that is

$$\chi_1 \leq \frac{2dJ}{\Delta}. \tag{33}$$

Define a polynomial function F_p of real variables $x_1, \ldots, x_{p-1}$ according to

$$F_p(x_1, \ldots, x_{p-1}) = x_{p-1} + \frac{1}{2} \sum_{p_1+p_2=p-1} x_{p_1} x_{p_2} + \frac{1}{6} \sum_{p_1+p_2+p_3=p-1} x_{p_1} x_{p_2} x_{p_3} + \frac{1}{24} \sum_{p_1+\ldots+p_4=p-1} x_{p_1} x_{p_2} x_{p_3} x_{p_4}. \tag{34}$$

Applying Lemma 3 and triangle inequality to Eq. (28) one gets

$$\chi_p \leq \frac{2^{13} dJ}{\Delta} F_p(\chi_1, \ldots, \chi_{p-1}), \quad p \geq 2. \tag{35}$$

To simplify notations, define constants

$$a = \frac{2dJ}{\Delta}, \quad b = \frac{2^{13} dJ}{\Delta}, \tag{36}$$

so that $\chi_1 \leq a$ and $\chi_p \leq bF_p(\chi_1, \ldots, \chi_{p-1})$ for $p \geq 2$. Consider the formal power series

$$\mu(\epsilon) = \sum_{p=1}^{\infty} \mu_p \epsilon^p, \quad \mu_1 = a, \quad \mu_p = bF_p(\mu_1, \ldots, \mu_{p-1}), \quad p \geq 2. \tag{37}$$

Since the polynomial F_p has non-negative coefficients one can prove inductively that $\chi_p \leq \mu_p$ for all $p \geq 1$. Hence it suffices to prove that the series Eq. (37) converges absolutely for $|\epsilon| \leq 2\epsilon_0$.

Our strategy will be to guess a function $\mu(\epsilon)$ analytic for $|\epsilon| \leq 2\epsilon_0$ whose Taylor series at $\epsilon = 0$ coincides with the series Eq. (37). By inspecting the recursive relation Eq. (37) one can easily convince oneself that $\mu(\epsilon)$ has to obey the following equation

$$\mu(\epsilon) = a\,\epsilon + b\,\epsilon \left(\mu(\epsilon) + \frac{1}{2}\mu^2(\epsilon) + \frac{1}{6}\mu^3(\epsilon) + \frac{1}{24}\mu^4(\epsilon) \right). \tag{38}$$

We can use it to write down the inverse function

$$\epsilon(\mu) = \frac{\mu}{Q(\mu)}, \quad Q(\mu) = a + b\left(\mu + \frac{1}{2}\mu^2 + \frac{1}{6}\mu^3 + \frac{1}{24}\mu^4\right). \tag{39}$$

Simple algebra shows that

$$|Q(\mu)| \geq \frac{a}{2} \quad \text{if} \quad |\mu| \leq \frac{a}{4b} = 2^{-14}. \tag{40}$$

Thus $\epsilon(\mu)$ is analytic for $|\mu| \leq 2^{-14}$. Define a set $M = \{\mu : |\mu| \leq 2^{-15}\}$.

Claim 2. *Let ϵ be a complex number such that $|\epsilon| \leq 2\epsilon_0$. Then equation $\epsilon(\mu) = \epsilon$ has a unique solution $\mu \in M$.*

13

Proof: One can easily show that for any $\mu_1, \mu_2 \in M$

$$|Q(\mu_1) - Q(\mu_2)| \leq 2b|\mu_1 - \mu_2|. \tag{41}$$

Assume $\epsilon(\mu_1) = \epsilon(\mu_2) = \epsilon$ for some $\mu_1, \mu_2 \in M$. If $\epsilon = 0$ then $\mu_1 = \mu_2 = 0$. Assume $\epsilon \neq 0$. Then $\mu_1, \mu_2 \neq 0$ and

$$\mu_1 - \mu_2 = \frac{\mu_2(Q(\mu_1) - Q(\mu_2))}{Q(\mu_2)}.$$

Applying the lower bound Eq. (40) and the upper bound Eq. (41) we get

$$|\mu_1 - \mu_2| \leq \frac{2^{-15}|Q(\mu_1) - Q(\mu_2)|}{a/2} \leq \frac{2^{-13} b|\mu_1 - \mu_2|}{a} \leq \frac{1}{2}|\mu_1 - \mu_2|.$$

Thus $\mu_1 = \mu_2$ and equation $\epsilon(\mu) = \epsilon$ has at most one solution $\mu \in M$. Therefore $\epsilon : M \to \epsilon(M)$ is an injection. Let us prove that $\epsilon(M)$ contains a ball of radius $2\epsilon_0$. Indeed, $\epsilon(M)$ is an open set and $0 \in \epsilon(M)$. Let γ be the boundary of M, i.e., a circle of radius 2^{-15} centered at 0. Then $\epsilon(\gamma)$ is the boundary of $\epsilon(M)$. For any $\mu \in \gamma$ one has $|Q(\mu)| \leq a + 2b|\mu| = a + 2^{-14}b \leq 2a$. Thus $\epsilon(M)$ contains a ball of radius

$$R = \min_{\mu \in \gamma} |\epsilon(\mu)| \geq \frac{2^{-15}}{2a} = \frac{2^{-17}\Delta}{dJ} = 2\epsilon_0$$

It completes the proof of the claim.

Let $K = \{\epsilon : |\epsilon| \leq 2\epsilon_0\}$. From Claim 2 we infer that $\epsilon(\mu)$ is an analytic bijection from the set $\epsilon^{-1}(K) \subseteq M$ to the set K. It follows from the inverse function theorem for analytic functions, see [10], that the inverse function $\mu(\epsilon)$ is analytic for $\epsilon \in K$. Therefore the series Eq. (37) converges absolutely for $|\epsilon| \leq 2\epsilon_0$.

The upper bound on μ_p can be obtained by standard methods using Cauchy's formula:

$$\mu_p = \frac{1}{2\pi i} \oint_{|\epsilon|=2\epsilon_0} \frac{\mu(\epsilon)d\epsilon}{\epsilon^{p+1}}.$$

Thus

$$|\mu_p| \leq \frac{1}{(2\epsilon_0)^p} \max_{\epsilon : |\epsilon|=2\epsilon_0} |\mu(\epsilon)| \leq \frac{2^{-15}}{(2\epsilon_0)^p}.$$

Recall that $\chi_p \leq \mu_p$, so the lemma is proved.

$\square$

One can summarize the results of this subsection as follows.

Corollary 3. *Suppose $|\epsilon| \leq 2\epsilon_0$. Then the Kirkwood-Thomas equations Eq. (17) have a unique solution C defined by the power series Eq. (25) with $\|C\|_1 \leq 2^{-15}$.*

3.3   Convergence of E-series

In this subsection we analyze convergence of the series $E(\epsilon) = \sum_{p=1}^{\infty} E_p \, \epsilon^p$ for the eigenvalue obtained from the Kirkwood-Thomas equation, see Eq. (4).

14

Lemma 6. *The series $E(\epsilon) = \sum_{p=1}^{\infty} E_p\,\epsilon^p$ converges absolutely for*

$$|\epsilon| \le 2\epsilon_0 = \frac{2^{-17}\Delta}{dJ}. \tag{42}$$

Besides,

$$|E_p| \le \frac{2^{-16}\,n\Delta}{(2\epsilon_0)^p} \quad for \quad p \ge 1. \tag{43}$$

Proof: Applying Lemma 3 and triangle inequality to Eq. (29) one gets

$$|E_1| \le ndJ, \quad |E_p| \le 2^4 ndJF_p(\chi_1,\dots,\chi_{p-1}), \quad p \ge 2, \tag{44}$$

where $\chi_p = \|C_p\|_1$ and the polynomial F_p is defined in Eq. (34). Define a formal series

$$e(\epsilon) = \sum_{p=1}^{\infty} e_p\,\epsilon^p, \quad e_1 = ndJ, \quad e_p = 2^4 ndJF_p(\chi_1,\dots,\chi_{p-1}), \quad p \ge 2. \tag{45}$$

By definition, $|E_p| \le e_p$ for all p. Besides, $e(\epsilon)$ can be expressed in terms of $\chi(\epsilon) = \sum_{p=1}^{\infty} \chi_p\,\epsilon^p$ as

$$e(\epsilon) = 2^4 ndJ\epsilon\left(\chi(\epsilon) + \frac{1}{2}\chi^2(\epsilon) + \frac{1}{6}\chi^3(\epsilon) + \frac{1}{24}\chi^4(\epsilon)\right) + ndJ\,\epsilon.$$

This equality can be verified by equating coefficients for each power of ϵ. Lemma 5 implies that $\chi(\epsilon)$ is analytic for $|\epsilon| \le 2\epsilon_0$. Therefore $e(\epsilon)$ and $E(\epsilon)$ are analytic for $|\epsilon| \le 2\epsilon_0$ and the first statement of the lemma is proved. In order to get an upper bound on e_p (and thus on E_p), use Cauchy's formula:

$$e_p = \frac{1}{2\pi i} \oint_{|\epsilon|=2\epsilon_0} \frac{e(\epsilon)d\epsilon}{\epsilon^{p+1}}.$$

It follows from Lemma 5 that $|\chi(\epsilon)| \le 2^{-15}$ for $|\epsilon| \le 2\epsilon_0$. Therefore

$$|e_p| \le \frac{1}{(2\epsilon_0)^p} \max_{\epsilon:|\epsilon|=2\epsilon_0} |e(\epsilon)| \le \frac{1}{(2\epsilon_0)^p}\left(2^4 ndJ(2\epsilon_0)2^{-14} + ndJ(2\epsilon_0)\right) \le \frac{2^{-16}n\Delta}{(2\epsilon_0)^p}.$$

The lemma is proved.

$\square$

Corollary 4. *Suppose $|\epsilon| \le 2\epsilon_0$. Then the series Eq. (29) converges absolutely to the smallest eigenvalue of $H(\epsilon)$. The smallest eigenvalue is non-degenerate and is separated from the rest of the spectrum by a gap at least $\Delta/2$.*

Proof: It follows from Corollary 2. Indeed, we have already shown that the conditions of Corollary 2 are satisfied with $\epsilon'_c = 2\epsilon_0$ and $c_{max} = 2^{-15}$, see Corollary 3. It yields $\epsilon''_c \le 2^{-15}\Delta/(dJ)$. Thus $\epsilon_c = \min(\epsilon'_c, \epsilon''_c) = 2\epsilon_0$.

$\square$

15

Lemma 6 allows one to estimate an error resulting from truncation of the series for the ground state energy at a finite order p.

Corollary 5. *Suppose $|\epsilon| \leq \epsilon_0$. Then*

$$\left| E(\epsilon) - n \sum_{q=1}^{p} E_q \, \epsilon^q \right| \leq n\Delta 2^{-16-p}. \tag{46}$$

Proof: Use Eq. (43).

$\square$

Summarizing, we have proved Theorems 1,2.

4 Linked cluster theorems

Throughout this section we shall use a term *linked cluster* which refers to a subset of vertices inducing a connected subgraph of $\mathcal{G}$. More formally,

Definition 2. *A subset $M \subseteq \mathcal{L}$ is called a linked cluster iff for any $u, v \in \mathcal{M}$ there exists a sequence of vertices $u_0, u_1, \ldots, u_t \in M$ such that $u_0 = u$, $u_t = v$ and $(u_j, u_{j+1}) \in \mathcal{E}$ for all $j = 0, \ldots, t-1$.*

Definition 3. *A connected size of a subset $M \subseteq \mathcal{L}$ is the minimal size of a linked cluster that contains all vertices of M. We shall denote a connected size of M as $|M|_c$.*

4.1 Linked cluster expansion for the ground state

Lemma 7. *Let $C(M) = \sum_{p=1}^{\infty} C_p(M)\epsilon^p$ be the solution of the Kirkwood-Thomas equations obtained in Section 3. Then*

$$C_p(M) = 0 \quad unless \quad |M|_c \leq p + 1. \tag{47}$$

Proof: We shall prove the lemma by induction in p. From Eq. (27) one infers that $C_1(M) = 0$ unless $M \subseteq \{u, v\}$ for some edge $(u, v) \in \mathcal{E}$. In particular, $C_1(M) = 0$ unless $|M|_c \leq 2$. It proves the statement of the lemma for $p = 1$. Suppose the statement is proved for $p = 1, \ldots, q-1$. From Eq. (28) one infers that $C_q(M)$ is a linear combination of terms like

$$x = C_{p_1}(M_1) \cdots C_{p_k}(M_k)\langle \Omega | a_M \hat{a}_{M_1}^{\dagger} \cdots \hat{a}_{M_k}^{\dagger}(V_{u,v})|\Omega\rangle, \tag{48}$$

where $p_1 + \ldots + p_k = q - 1$. Let us figure out under what circumstances the matrix element in Eq. (48) can be non-zero.

Claim 3. *Let $M, M_1, \ldots, M_k \subseteq \mathcal{L}$ be non-empty sets, $N = M_1 \cup \ldots \cup M_k$, $(u, v) \in \mathcal{E}$. Denote*

$$y = \langle \Omega | a_M \, \hat{a}_{M_1}^{\dagger} \cdots \hat{a}_{M_k}^{\dagger}(V_{u,v})|\Omega\rangle.$$

Then $y = 0$ unless the following conditions are met:
(i) Each set $M_1, \ldots, M_k$ contains at least one of the vertices u, v.
(ii) $N \backslash \{u, v\} \subseteq M \subseteq N \cup \{u, v\}$.

16

Remark: this claim is true even for $M = \emptyset$ if one adopts a convention $a_\emptyset = I$.

Proof: Suppose some set M_j contains neither u nor v. Then $a^\dagger_{M_j}$ commutes with $V_{u,v}$ as well as with all operators $a^\dagger_{M_i}$ for $i \neq j$. Thus $y = 0$. Suppose condition $N \backslash \{u, v\} \subseteq M$ is violated. Then there exists a set M_j and a vertex $w \in M_j$ such that $w \neq u, v$ and $w \notin M$. Thus $a^\dagger_{M_j}$ contains a factor $a^\dagger_w$ which commutes with all other operators involved in y. By moving $a^\dagger_w$ leftwards one can show that each of 2^k terms in y starts from $\langle \Omega | a^\dagger_w$, that is, $y = 0$. Suppose condition $M \subseteq N \cup \{u, v\}$ is violated, that is, there exists a vertex $w \in M$ such that $w \notin N$ and $w \neq u, v$. Then the operator a_M contains a factor a_w which commutes with all other operators involved in y. By moving a_w rightwards one can show that each of 2^k terms in y tails with $a_w |\Omega\rangle$, that is, $y = 0$.

$\square$

Returning to Eq. (48) we conclude that $x = 0$ unless each set M_j contains either u or/and v, and $M \subseteq N \cup \{u, v\}$. Let $\tilde{M}_j$ be a linked cluster of minimal size containing M_j, that is, $|M_j|_c = |\tilde{M}_j|$. Let $\tilde{N} = \tilde{M}_1 \cup \ldots \cup \tilde{M}_k$ and $C = \tilde{N} \cup \{u\} \cup \{v\}$. Then C is a linked cluster and $M \subseteq C$. Note that

$$|C| \leq \sum_{j=1}^k |\tilde{M}_j| + 2 - k = \sum_{j=1}^k |M_j|_c + 2 - k$$

where we have taken into account that each $\tilde{M}_j$ contains either u or/and v. By induction hypothesis we have $C_{p_j}(M_j) = 0$ unless $|M_j|_c \leq p_j + 1$. Thus for any non-zero term x one has

$$|C| \leq \sum_{j=1}^k (p_j + 1) + 2 - k = q - 1 + k + 2 - k = q + 1.$$

Thus $C_q(M) = 0$ unless $|M|_c \leq q + 1$. $\square$

4.2 Upper bound on the number of linked clusters

The following lemma asserts that the number of linked clusters of size p containing a given vertex grows at most exponentially with p (if the maximal degree of the graph d is a constant). To the best of our knowledge, this lemma has been originally proved in Ref. [21] by Aliferis, Gottesman, and Preskill in the context of fault-tolerant quantum computation[3].

Lemma 8. *Let $N_p(u)$ be the number of linked clusters with p vertices that contain a vertex u and $N_p = \max_{u \in \mathcal{L}} N_p(u)$. Then*

$$N_p \leq (4d)^{p-1}. \tag{49}$$

Proof: Let $\mathcal{T}_p(u)$ be a set of trees with p vertices that contain a vertex u (naturally, we consider only those trees that are subgraphs of $\mathcal{G}$). Let $T_p(u) = |\mathcal{T}_p(u)|$ be the number of such trees. For any tree $T \in \mathcal{T}_p(u)$, a set of vertices of T is a linked cluster that contains u. Conversely, if $M \ni u$ is a linked cluster, $|M| = p$, consider a subgraph G_M induced by M. Then any spanning tree of G_M belongs to $\mathcal{T}_p(u)$. Thus $N_p(u) \leq T_p(u)$.

---

[3]The authors became aware of it after completion of the present work.

Denote $T_p = \max_{u \in \mathcal{L}} T_p(u)$. Obviously, $T_1 = 1$ and $T_2 \leq d$. Let us prove that

$$T_p \leq d \sum_{p_1 + p_2 = p} T_{p_1} T_{p_2}, \tag{50}$$

where the convention $T_0 = 0$ is adopted. Indeed, for any edge e incident to a vertex u define a set $\mathcal{T}_p(u, e)$ that includes all trees $T \in \mathcal{T}_p(u)$ that contain an edge e. Let $T_p(u, e) = |\mathcal{T}_p(u, e)|$. Clearly,

$$\mathcal{T}_p(u) = \cup_e \mathcal{T}_p(u, e), \quad T_p(u) \leq \sum_e T_p(u, e) \leq d \max_e T_p(u, e). \tag{51}$$

Let $e = (u, v)$ be the edge that achieves the maximum. Note that any tree $T \in \mathcal{T}_p(u, e)$ consists of the edge (u, v) and two disjoint trees $T_1 \in \mathcal{T}_{p_1}(u)$ and $T_2 \in \mathcal{T}_{p_2}(v)$, where $p_1 + p_2 = p$. Thus we have an upper bound

$$T_p(u, e) \leq \sum_{p_1 + p_2 = p} T_{p_1}(u) T_{p_2}(v) \leq \sum_{p_1 + p_2 = p} T_{p_1} T_{p_2}.$$

Substituting it into Eq. (51) and taking the maximum over $u \in \mathcal{L}$ we obtain Eq. (50).

Define a sequence $S_1, S_2, \ldots$ such that

$$S_1 = 1, \quad S_p = d \sum_{p_1 + p_2 = p} S_{p_1} S_{p_2} \quad \text{for} \quad p \geq 2. \tag{52}$$

Clearly, $T_1 = S_1 = 1$ and $T_2 \leq d = S_2$. It follows that $T_p \leq S_p$ for all p. In order to derive an explicit formula for S_p define a generating function $S(x) = \sum_{p=1}^{\infty} S_p x^p$. It obeys an equation $S(x) = dS(x)^2 + x$, which implies

$$S(x) = \frac{1 - \sqrt{1 - 4dx}}{2d}.$$

Taking the derivatives one gets

$$S_p = \frac{1}{p!} \frac{d^p S}{dx^p}\bigg|_{x=0} = -\frac{(4d)^p}{2d(p!)} \prod_{a=0}^{p-1} \left(a - \frac{1}{2}\right).$$

It follows that

$$S_p \leq \frac{(4d)^p (p-1)!}{4d(p!)} \leq \frac{(4d)^{p-1}}{p} \leq (4d)^{p-1}.$$

Summarizing, $N_p \leq T_p \leq S_p \leq (4d)^{p-1}$.

$\square$

4.3 Linked cluster expansion for the ground state energy

This subsection provides the necessary tools for computing spin-spin correlators. A reader interested only in computing the ground state energy can safely skip it.

Let us consider a more general family of Hamiltonians for which the parameter ϵ may be different on different edges. Let variable $\epsilon_{u,v}$ be assigned to an edge $(u, v) \in \mathcal{E}$. For any subset of edges $A \subseteq \mathcal{E}$ denote $\epsilon[A]$ a collection of variables assigned to edges of A. The Hamiltonian is

$$H(\epsilon[\mathcal{E}]) = H_0 + \sum_{(u,v) \in \mathcal{E}} \epsilon_{u,v} V_{u,v}. \tag{53}$$

Let $E(\epsilon[\mathcal{E}])$ be the ground state energy of $H(\epsilon[\mathcal{E}])$. We shall consider multivariate Taylor series for the function $E(\epsilon[\mathcal{E}])$.

Lemma 9. *The multivariate Taylor series for $E(\epsilon[\mathcal{E}])$ at the point $\epsilon[\mathcal{E}] = 0$ converges absolutely if $|\epsilon_{u,v}| \leq \epsilon_0$ for all $(u,v) \in \mathcal{E}$.*

Proof: Let $\Omega = \{\epsilon[\mathcal{E}] \ : \ |\epsilon_{u,v}| \leq 2\epsilon_0$ for all $(u,v) \in \mathcal{E}\}$. Let us firstly show $E(\epsilon[\mathcal{E}])$ is analytic function of each individual variable $\epsilon_{u,v}$ for $\epsilon[\mathcal{E}] \in \Omega$. Indeed, let $\tilde{\mathcal{E}}$ be a set of all edges except (u,v). Define an unperturbed Hamiltonian $\tilde{H}_0 = H_0 + \sum_{(u,v)\in\tilde{\mathcal{E}}} \epsilon_{u,v} V_{u,v}$ and a perturbation $\epsilon_{u,v} V_{u,v}$. It follows from Corollary 4 that $\tilde{H}_0$ has non-degenerate ground state and the spectral gap at least $\Delta/2$. Applying the standard perturbation theory to a perturbed Hamiltonian $\tilde{H}_0 + \epsilon_{u,v} V_{u,v}$ we conclude that $E(\epsilon[\mathcal{E}])$ is analytic function of $\epsilon_{u,v}$ as long as the Weyl condition $\|\epsilon_{u,v} V_{u,v}\| < \Delta/4$ is satisfied. Since we assumed that $|\epsilon_{u,v}| \leq 2\epsilon_0$, one has $\|\epsilon_{u,v} V_{u,v}\| < 2\epsilon_0 J = 2^{-17}\Delta/d < \Delta/4$. Thus $E(\epsilon[\mathcal{E}])$ is analytic in Ω with respect to each individual variable $\epsilon_{u,v}$. Repeatedly using Cauchy's formula one gets

$$E(\epsilon[\mathcal{E}]) = \left(\prod_{(u,v)\in\mathcal{E}} \frac{1}{2\pi i} \oint_{|z_{u,v}|=2\epsilon_0} \frac{1}{(z_{u,v} - \epsilon_{u,v})} \right) E(z[\mathcal{E}]). \tag{54}$$

Since H_0 and V are bounded operators, the absolute value $|E(z[\mathcal{E}])|$ can be bounded by a constant (maybe depending on n). The Taylor series in $\epsilon_{u,v}$ at the point $\epsilon_{u,v} = 0$ for any factor $1/(z_{u,v} - \epsilon_{u,v})$ in Eq. (54) converges absolutely as long as $|\epsilon_{u,v}| < 2\epsilon_0$. Thus the Taylor series for $E(\epsilon[\mathcal{E}])$ converges absolutely if $|\epsilon_{u,v}| \leq \epsilon_0$ for all $(u,v) \in \mathcal{E}$.

□

The Taylor series for $E(\epsilon[\mathcal{E}])$ can be uniquely written in the form

$$E(\epsilon[\mathcal{E}]) = \sum_{A\subseteq\mathcal{E}} \left(\prod_{(u,v)\in A} \epsilon_{u,v} \right) p_A(\epsilon[A]), \tag{55}$$

where the sum is over all subsets of edges A and $p_A(\epsilon[A])$ is the series that involves only variables $\epsilon_{u,v}$ pertaining to A. Clearly, the coefficients of $p_A(\epsilon[A])$ are functionals of interactions $V_{u,v}$ with $(u,v) \in A$ only. The main goal of this section is to show that the expansion Eq. (55) involves only linked clusters of edges. Let us firstly define this notion.

Definition 4. *A subset of edges $A \subseteq \mathcal{E}$ is called a linked cluster iff the subset of vertices induced by A is a linked cluster.*

Lemma 10. *The series Eq. (55) involves only linked clusters of edges A.*

Proof: Suppose $A \subseteq \mathcal{E}$ is not a linked cluster of edges. Let $M \subseteq \mathcal{L}$ be a set of vertices induced by A. Since M is not a linked cluster, it can be represented as a disjoint union $M = M_1 \cup M_2$, where $M_1, M_2 \subseteq \mathcal{L}$, $M_1 \cap M_2 = \emptyset$, and no edge connects M_1 and M_2. Accordingly, A can be represented as a union $A = A_1 \cup A_2$, where A_1 and A_2 are the set of edges inducing M_1 and M_2 respectively. Let us choose variables $\epsilon[\mathcal{E}]$ such that $\epsilon_{u,v} = 0$ unless $(u,v) \in A$. Then it is clear that the Hamiltonian $H(\epsilon[\mathcal{E}])$ splits into a sum of three terms acting on non-overlapping sets of qubits:

$$H(\epsilon[\mathcal{E}]) = H_1 + H_2 + H_{else}, \quad H_j = \sum_{u\in M_j} \Delta_u |1\rangle\langle 1|_u + \sum_{(u,v)\in A_j} \epsilon_{u,v} V_{u,v}, \quad H_{else} = \sum_{u\in\mathcal{L}\setminus(M_1\cup M_2)} \Delta_u |1\rangle\langle 1|_u.$$

19

The ground state energy of $H(\epsilon[\mathcal{E}])$ is equal to the sum of ground state energies of H_1, H_2, and H_{else}. It implies that

$$E(\epsilon[\mathcal{E}]) = E(\epsilon[A_1]) + E(\epsilon[A_2]). \tag{56}$$

If we assume that $p_A(\epsilon[A]) \neq 0$, when $E(\epsilon[\mathcal{E}])$ would include at least one monomial including variables from both sets A_1, A_2 which contradicts to Eq. (56).

$\square$

The following implication of Lemma 10 will simplify computation of spin-spin correlators .

Corollary 6. *Consider a Hamiltonian $H = H_0 + \epsilon V$, where $V = \sum_{(u,v) \in \mathcal{E}} V_{u,v}$. Let $E(\epsilon) = \sum_{p=1}^{\infty} E_p \, \epsilon^p$ be the series for the ground state energy of H. Suppose the interaction $V_{s,t}$ depends on a parameter η for some edge $(s,t) \in \mathcal{E}$. Then a derivative*

$$K_p = \left. \frac{\partial E_p}{\partial \eta} \right|_{\eta=0}$$

can be computed by setting $V_{u,v} = 0$ for all edges (u,v) having distance $p+1$ or greater from the edge (s,t).

Proof: Indeed, E_p can be obtained from Eq. (55) by setting $\epsilon_{u,v} = \epsilon$ on all edges, restricting the sum to linked clusters A of size at most p and collecting all monomials of total degree p. Clusters A that do not contain the edge (s,t) will not contribute to K_p. Clusters A that contain the edge (s,t) cannot contain any edge (u,v) having distance $p+1$ or greater from the edge (s,t).

$\square$

5 Computational algorithms

In this section we describe an algorithm that takes as input a description of the Hamiltonians H_0, V and an integer p. The algorithm returns a list of coefficients $E_1, \ldots, E_p$ in the series for the ground state energy $E(\epsilon) = \sum_{p=1}^{\infty} E_p \, \epsilon^p$. The running time of the algorithm is $n \exp(O(p))$. In Section 5.3 we describe a generalization of the algorithm that allows one to compute spin-spin correlation functions.

5.1 Computing the coefficients $C_p(M)$

The first part of the algorithm is to compute the coefficients $C_q(M)$, $\emptyset \neq M \subseteq \mathcal{L}$ sequentially for $q = 1, \ldots, p$ using the solution of the Kirkwood-Thomas equations, see Eqs. (27,28). This gives an approximate description of the ground state.

We shall store triples $(M, q, C_q(M))$ in n bins (memory registers) B_u labeled by vertices of the graph $u \in \mathcal{L}$. Once a coefficient $C_q(M)$ is computed, the triple $(M, q, C_q(M))$ is placed into every bin B_u for which $u \in M$. From Lemma 7 we learn that $C_q(M) = 0$ unless M is a subset of some linked cluster $\tilde{M}$ of size at most $q + 1$. According to Lemma 8, the number of linked clusters $\tilde{M}$ of size $q + 1$ containing vertex u is bounded by $\exp(O(q))$, where the coefficient in the exponent depends only on d. Each linked cluster of size $q + 1$ containing vertex u has 2^q subsets containing

20

vertex u. Thus we can bound the number of entries in the bin B_u at the moment when all coefficients $C_1(M), \ldots, C_p(M)$ have been computed as $|B_u| \leq \sum_{q=1}^{p} 2^q \exp(O(q)) = \exp(O(p))$.

Suppose we have already computed all non-zero coefficients $C_1(M), \ldots, C_{q-1}(M)$, $M \subseteq \mathcal{L}$. The next step is to compute coefficients $C_q(M)$ for all sets $M \subseteq \mathcal{L}$ satisfying the condition of Lemma 7, that is $|M|_c \leq q + 1$. Expanding Eq. (28) one gets

$$C_q(M) = E_0(M)^{-1} \sum_{(u,v) \in \mathcal{E}} \sum_{k=1}^{4} \frac{1}{k!} \sum_{p_1 + \ldots + p_k = q-1} \sum_{M_1, \ldots, M_k \subseteq \mathcal{L}} C_{p_1}(M_1) \cdots C_{p_k}(M_k) \langle \Omega | a_M \hat{a}_{M_1} \cdots \hat{a}_{M_k}(V_{u,v}) | \Omega \rangle. \tag{57}$$

Note that the right hand side of this equation involves only coefficients $C_{p_j}(M_j)$ that have been already computed. For simplicity let us assume that computation of any term in Eq. (57) requires one unit of time[4]. Denote

$$x = \langle \Omega | a_M \, \hat{a}^\dagger_{M_1} \cdots \hat{a}^\dagger_{M_k}(V_{u,v}) | \Omega \rangle. \tag{58}$$

Recall, see Claim 3, that $x = 0$ unless the following conditions are met:
(i) Each set $M_1, \ldots, M_k$ contains at least one of the vertices u, v.
(ii) $N \backslash \{u, v\} \subseteq M \subseteq N \cup \{u, v\}$, where $N = M_1 \cup \ldots \cup M_k$.
The property (i) implies that for a fixed edge (u, v) we can restrict the three rightmost sums in Eq. (57) by taking triples $(M_j, p_j, C_{p_j}(M_j))$ either from the bin B_u or from the bin B_v. Therefore for a fixed (u, v), the overall number of non-zero terms in the three rightmost sums in Eq. (57) can be bounded by $(|B_u| + |B_v|)^k \leq (|B_u| + |B_v|)^4 = \exp(O(q))$.

We shall now prove that only a small number of edges (u, v) can give a non-zero contribution to $C_q(M)$. Indeed, there are two cases: (1) $M \subseteq \{u, v\}$; (2) There exists $w \in M$ such that $w \notin \{u, v\}$. Clearly only $O(1)$ edges (u, v) can lead to the case (1), so let us focus on the case (2). Consider any term x as in Eq. (58). Properties (i),(ii) above imply that $x = 0$ unless there exists a set M_j such that $w \in M_j$ and one of the vertices u, v belongs to M_j. Without loss of generality, $w, u \in M_j$. Lemma 7 implies that $C_{p_j}(M_j) = 0$ unless $|M_j|_c \leq p_j + 1 \leq q$. Therefore the distance between u and w is at most q. Taking into account that $|M| \leq |M|_c \leq q + 1$, we can bound the number of edges (u, v) that can give a non-zero contribution to $C_q(M)$ by $|M| d^{q+1} \leq (q+1) d^{q+1} = \exp(O(q))$. Summarizing, the overall number of non-zero terms in Eq. (57) is $\exp(O(q))$.

In order to compute all non-zero coefficients $C_q(M)$ we will have to repeat the procedure above for each subset M satisfying the condition of Lemma 7, that is $|M|_c \leq q + 1$. By Lemma 8, the number of such sets is $n \exp(O(q))$. Summarizing, the overall time one needs to compute all the coefficients $C_1(M), \ldots, C_p(M)$ is $n \exp(O(p))$.

5.2 Computing the ground state energy

The final step of the algorithm is to compute the coefficients $E_1, \ldots, E_p$ using Eq. (29). This equation can be expanded as

$$E_p = \sum_{(u,v) \in \mathcal{E}} \sum_{k=1}^{4} \frac{1}{k!} \sum_{p_1 + \ldots + p_k = p-1} \sum_{M_1, \ldots, M_k \subseteq \mathcal{L}} C_{p_1}(M_1) \cdots C_{p_k}(M_k) \langle \Omega | \hat{a}_{M_1} \cdots \hat{a}_{M_k}(V_{u,v}) | \Omega \rangle. \tag{59}$$

---

[4] This assumption might seem unjustified, because the precision up to which the coefficient $C_q(M)$ must be computed depends upon δ. However, taking into account these subtleties will lead to an additional overhead $poly(\log n, \log \delta^{-1})$ which can be neglected since the algorithm has running time $poly(n, \delta^{-1})$.

Expanding the commutators in the matrix element one gets 2^k terms. However, only the term in which $V_{u,v}$ is the leftmost operator gives a non-zero contribution. Thus

$$\langle \Omega | \hat{a}_{M_1} \cdots \hat{a}_{M_k} (V_{u,v}) | \Omega \rangle = (-1)^k \langle \Omega | V_{u,v}\, a^\dagger_{M_k} \cdots a^\dagger_{M_1} | \Omega \rangle.$$

In particular, we can restrict the summation over $M_1, \ldots, M_k$ by subsets of $\{u, v\}$ only. There are only 3 such subsets: $\{u\}$, $\{v\}$, and $\{u, v\}$. This observation implies that the number of terms in the rightmost sum in Eq. (59) is $O(1)$. Since there are $O(p^3)$ partitions $p_1 + \ldots + p_k = p - 1$, $k \le 4$, the overall number of terms in Eq. (59) is $O(np^3)$.

Combining the results of Subsections 5.1, 5.2 we conclude that the overall time needed to compute the coefficients $E_1, \ldots, E_p$ scales as $n \exp{(O(p))}$. In the above analysis we did not keep track of the coefficients in the exponents $O(p)$. If one computes the exact coefficient, it yields the overall running time $n2^{15+6\log{(d)}}$, where log stands for the base two logarithm. Accordingly, the running time as a function of n and δ scales as

$$T(n, \delta) \sim n(n\delta^{-1})^{15+6\log{(d)}}.$$

For example, implementing the algorithm on a 2D square lattice $(d = 4)$ would require a running time $T(n, \delta) \sim n(n\delta^{-1})^{27}$, which is certainly not practical. Note however, that the power of $n\delta^{-1}$ depends upon the ratio $|\epsilon|/R$, where R is the convergence radius of the series Eq. (4). The power $15 + 6\log{(d)}$ corresponds to the most pessimistic scenario $R = 2\epsilon_0$ (the best lower bound on the convergence radius that we can prove) and $|\epsilon| = \epsilon_0$.

5.3 Computing spin-spin correlation functions

Let $s, t \in \mathcal{L}$ be any pair of vertices. It may or may not be an edge of the graph $\mathcal{G}$. Let us add (s, t) to the set of edges $\mathcal{E}$ (by creating a double edge between s and t if necessary). The modified graph has maximal degree $d^* = d + 1$. Let $O_{s,t}$ be a Hermitian operator acting non-trivially only on qubits s, t. We shall assume that $\|O_{s,t}\| \le J$. The quantity we are interested in is the expectation value

$$K = \frac{\langle \psi | O_{s,t} | \psi \rangle}{\langle \psi | \psi \rangle},$$

where $|\psi\rangle$ is the ground state of $H(\epsilon) = H_0 + \epsilon V$. Our goal is to compute K with a specified precision δ. To this end let us define a Hamiltonian

$$H(\epsilon, \eta) = H_0 + \epsilon V + \eta O_{s,t}. \tag{60}$$

Let $E(\epsilon, \eta)$ be the smallest eigenvalue of $H(\epsilon, \eta)$. As we know from Lemma 9, the Taylor series

$$E(\epsilon, \eta) = \sum_{p,q=0}^{\infty} E_{p,q}\, \epsilon^p\, \eta^q \tag{61}$$

converges absolutely for $|\epsilon|, |\eta| \le \epsilon_0^*$, where

$$\epsilon_0^* = \frac{2^{-18}\,\Delta}{d^*J} = \frac{2^{-18}\,\Delta}{(d+1)J}.$$

22

The Hellman-Feynman theorem asserts that

$$K = \left. \frac{\partial E(\epsilon, \eta)}{\partial \eta} \right|_{\eta=0} = \sum_{p=0}^{\infty} E_{p,1} \, \epsilon^p. \tag{62}$$

Our algorithm will get an approximation to K by computing a truncation of series in Eq. (62). The following lemma provides a bound on the error resulting from the truncation.

Lemma 11. *Suppose $|\epsilon| \leq \epsilon_0^*/(2d)$. Then*

$$\left| K - \sum_{q=0}^{p} E_{q,1} \, \epsilon^q \right| \leq 2^{-16-p} \, Jd(d+1). \tag{63}$$

Proof: Let us firstly prove that

$$|E_{p,1}| \leq \frac{2^{-16} \, d^{p+1} \, \Delta}{(\epsilon_0^*)^{p+1}}. \tag{64}$$

Indeed, use Cauchy's formula:

$$E_{p,1} = \frac{1}{(2\pi i)^2} \oint_{|\epsilon|=\epsilon_0^*} \oint_{|\eta|=\epsilon_0^*} \frac{E(\epsilon, \eta) \, d\epsilon \, d\eta}{\epsilon^{p+1} \, \eta^2}. \tag{65}$$

From Lemma 6 we infer that $|E(\epsilon, \eta)| \leq 2^{-16} \, n\Delta$. However, we would like to have an upper bound independent of n. To this end we employ Corollary 6 according to which $E_{p,1}$ can be computed by restricting the Hamiltonian on $(d+1)$-neighborhood of the edge (s,t). The number of spins in this neighborhood is at most $n^* = d^{p+1}$. Therefore $|E(\epsilon, \eta)| \leq 2^{-16} \, d^{p+1}\Delta$. Substituting this bound into Eq. (65) one gets Eq. (64). Finally, using the condition $|\epsilon| \leq \epsilon_0^*/(2d)$ we bound the sum $\sum_{q=p+1}^{\infty} |E_{q,1}| \, \epsilon^q$ as in Eq. (63).

$\square$

Lemma 11 shows that in order to compute K with an absolute error δ it is enough to compute the coefficients $E_{0,1}, E_{1,1}, \ldots, E_{p,1}$ in the series Eq. (61) with $p = \log(\delta^{-1}) + O(1)$.

Computation of the coefficients $E_{p,1}$ requires only minor modifications of the algorithm described in Sections 5.1,5.2. Indeed, consider a function $\tilde{E}(\epsilon, \eta) = E(\epsilon, \epsilon\eta)$. Using the series Eq. (61) one gets

$$\tilde{E}(\epsilon, \eta) = \sum_{r=1}^{\infty} \tilde{E}_r(\eta) \, \epsilon^r, \quad \tilde{E}_r(\eta) = \sum_{p+q=r} E_{p,q} \, \eta^q.$$

In particular,

$$E_{p,1} = \left. \frac{\partial \tilde{E}_{p+1}(\eta)}{\partial \eta} \right|_{\eta=0}. \tag{66}$$

On the other hand, $\tilde{E}(\epsilon, \eta)$ is the ground state energy of a Hamiltonian $H_0 + \epsilon \, (V + \eta \, O_{s,t})$. Thus we can compute the coefficients $\tilde{E}_1(\eta), \ldots, \tilde{E}_{p+1}(\eta)$ using the already available algorithm for the ground state energy. Moreover, from Corollary 6 we know that the coefficients $E_{0,1}, E_{1,1}, \ldots, E_{p,1}$ can be computed by restricting the Hamiltonian to the $(p+1)$-neighborhood of the edge (s,t). Thus we can apply Theorem 3 with n replaced by $n^* = d^{p+1}$, obtaining an algorithm with a running time $\exp(O(p))$ for computing $\tilde{E}_1(\eta), \ldots, \tilde{E}_{p+1}(\eta)$. In fact, at every step of this algorithm we have to retain only the terms independent of η and the terms linear in η, see Eq. (66). Since we have chosen $p = \log(\delta^{-1}) + O(1)$, the running time of the algorithm is $poly(\delta^{-1})$.

6 Discussion and open problems

We have proved that the ground state properties of a spin Hamiltonian with sufficiently weak interactions between qubits can be computed efficiently. We hope that this result could be generalized in several different directions. Firstly, one could try to consider more general class of unperturbed Hamiltonians H_0, for example, classical Ising-like Hamiltonians. In addition, one could consider systems of fermionic modes rather than spins. Secondly, one could investigate possible generalizations of the Kirkwood-Thomas ansatz to the case of degenerate ground state. In this case the ansatz should be constructed for an effective Hamiltonian acting on a low-energy subspace rather than for the ground state. Results of this kind could provide a rigorous basis for perturbative derivations of low-energy effective Hamiltonians, for example the mapping from the half-filled Hubbard model to the Heisenberg model. Thirdly, one could try to get a stronger lower bound on the convergence radius R of the series $E(\epsilon) = \sum_{p=1}^{\infty} E_p \, \epsilon^p$. We note that a stronger lower bound $R \geq \Delta/dJ$ can be easily obtained for classical Hamiltonians, when all interactions $V_{u,v}$ are diagonal in the $|0\rangle, |1\rangle$ basis. Therefore, one could speculate that in the quantum case R should be close to Δ/dJ.

7 Acknowledgments

8 Appendix A

In this section we prove Lemma 3. By definition of the norm, $\|C\|_1 = \max_{u \in \mathcal{L}} Y_u$, where

$$Y_u = \sum_{M \ni u} E_0(M)^{-1} \left| \langle \Omega | a_M \, \hat{C}_1 \cdots \hat{C}_k(V) | \Omega \rangle \right|.$$

Applying the triangle inequality one can bound Y_u as

$$Y_u \;\; \leq \;\; X_u := \sum_{M \ni u} E_0(M)^{-1} \sum_{(v,w) \in \mathcal{E}} \sum_{M_1, \ldots, M_k} \left| \langle \Omega | a_M \, \hat{a}^\dagger_{M_1} \cdots \hat{a}^\dagger_{M_k} (V_{v,w}) | \Omega \rangle \right| \, |C_1(M_1)| \cdots |C_k(M_k)| \quad (67)$$

Here the last sum is over all non-empty subsets $M_1, \ldots, M_k \subseteq \mathcal{L}$. Claim 3 allows one to restrict the summation in Eq. (67) only by tuples $(M, M_1, \ldots, M_k, v, w)$ satisfying conditions (i),(ii). We shall partition X_u into $k + 1$ (possibly overlapping) sums that will be dealt with separately. We define $X_u^{(j)}$, $j = 1, \ldots, k$ as a sum of all terms in Eq. (67) for which $u \in M_j$. We define $X_u^{(0)}$ as a sum of all terms in Eq. (67) for which $u \in \{v, w\}$. In other words,

$$X_u^{(j)} = \sum_{M \ni u} E_0(M)^{-1} \sum_{(v,w) \in \mathcal{E}} \sum_{M_1, \ldots, M_k} \chi_{M_j}(u) \left| \langle \Omega | a_M \, \hat{a}^\dagger_{M_1} \cdots \hat{a}^\dagger_{M_k} (V_{v,w}) | \Omega \rangle \right| \, |C_1(M_1)| \cdots |C_k(M_k)|,$$

where χ_{M_j} is the characteristic function[5] of M_j and

$$X_u^{(0)} = \sum_{M \ni u} E_0(M)^{-1} \sum_{v \,:\, (u,v) \in \mathcal{E}} \sum_{M_1,\ldots,M_k} \left| \langle \Omega | a_M \, \hat{a}_{M_1}^\dagger \cdots \hat{a}_{M_k}^\dagger (V_{u,v}) | \Omega \rangle \right| |C_1(M_1)| \cdots |C_k(M_k)|$$

Condition (ii) in Claim 3 implies that $u \in M \subseteq N \cup \{v, w\}$, so that each term in Eq. (67) appears at least one time in the sums $X_u^{(0)}, \ldots, X_u^{(k)}$, hence

$$X_u \le \sum_{j=0}^{k} X_u^{(j)}. \tag{68}$$

Upper bound on $X^{(j)}$, $1 \le j \le k$: The property (i) in Claim 3 implies that at least one end-point of the edge (v, w) belongs to M_j. W.l.o.g. $v \in M_j$. Then property (ii) implies $M_j \subseteq M \cup \{w\}$, so that $|M_j| \le 2|M|$ (recall that M is a non-empty set because $u \in M$). It gives us a bound $E_0(M) \ge \Delta|M| \ge (\Delta/2)|M_j|$. Note also that for any fixed $M_1, \ldots, M_k$ and v, w there exist at most four sets M satisfying condition (ii) of Claim 3 (take N and add/subtract vertices v and w). Therefore

$$X_u^{(j)} \le \frac{8}{\Delta} \max_M \sum_{(v,w) \in \mathcal{E}} \sum_{M_1,\ldots,M_k} \chi_{M_j}(u) \chi_{M_j}(v) \frac{1}{|M_j|} \left| \langle \Omega | a_M \, \hat{a}_{M_1}^\dagger \cdots \hat{a}_{M_k}^\dagger (V_{v,w}) | \Omega \rangle \right| |C_1(M_1)| \cdots |C_k(M_k)|.$$

Now we can bound the matrix element by $2^k J$ and add a restriction $M_i \cap \{v, w\} \ne \emptyset$ to the summations over sets M_i, $i \ne j$, see Claim 3, property (i). Taking into account that

$$\sum_{M_i \,:\, M_i \cap \{v,w\} \ne \emptyset} |C_i(M_i)| \le \sum_{M_i \ni v} |C_i(M_i)| + \sum_{M_i \ni w} |C_i(M_i)| \le 2\|C_i\|_1 \tag{69}$$

we arrive to

$$X_u^{(j)} \le \frac{2^{2k+2} J}{\Delta} \prod_{i \ne j} \|C_i\|_1 \sum_{(v,w) \in \mathcal{E}} \sum_{M_j} \chi_{M_j}(u) \chi_{M_j}(v) \frac{1}{|M_j|} |C_j(M_j)|.$$

Changing the order of summations and bounding the sum over (v, w) by $d|M_j|$ one gets

$$X_u^{(j)} \le \frac{2^{2k+2} dJ}{\Delta} \prod_{i \ne j} \|C_i\|_1 \sum_{M_j} \chi_{M_j}(u) |C_j(M_j)| \le \frac{2^{2k+2} dJ}{\Delta} \prod_{i=1}^{k} \|C_i\|_1.$$

Finally, Lemma 2 implies that it is enough to consider $k \le 4$, so that

$$\sum_{j=1}^{k} X_u^{(j)} \le \frac{2^{12} dJ}{\Delta} \prod_{j=1}^{k} \|C_j\|_1. \tag{70}$$

Upper bound on $X^{(0)}$: Claim 3 implies that for any fixed $(M_1, \ldots, M_k, v)$ there exist at most four sets M satisfying (ii). Using a bound $E_0(M) \ge \Delta$ we arrive to

$$X_u^{(0)} \le \frac{4}{\Delta} \max_M \sum_{v \,:\, (u,v) \in \mathcal{E}} \sum_{M_1,\ldots,M_k} \left| \langle \Omega | a_M \, \hat{a}_{M_1}^\dagger \cdots \hat{a}_{M_k}^\dagger (V_{u,v}) | \Omega \rangle \right| |C_1(M_1)| \cdots |C_k(M_k)|.$$

---

[5]That is $\chi_{M_j}(u) = 1$ if $u \in M_j$ and $\chi_{M_j}(u) = 0$ otherwise.

Claim 3, property (i) allows us to bound the matrix element by $2^k J$ and add a restriction $M_i \cap \{u, v\} \neq \emptyset$ to the summations over sets M_i. Using Eq. (69) we arrive to

$$X_u^{(0)} \leq \frac{2^{2k+2}J}{\Delta} \prod_{j=1}^{k} \|C_j\|_1 \sum_{v\,:\,(u,v)\in\mathcal{E}} 1 \leq \frac{2^{10}dJ}{\Delta} \prod_{j=1}^{k} \|C_j\|_1. \tag{71}$$

Combining Eqs. (68,70,71) we prove the upper bound Eq. (19).

The second bound Eq. (20) of Lemma 3 is much easier to prove. Applying the triangle inequality one gets

$$|\langle\Omega|\hat{C}_1\cdots\hat{C}_k(V_{u,v})|\Omega\rangle| \leq \sum_{M_1,\ldots,M_k} |\langle\Omega|\hat{a}^\dagger_{M_1}\cdots\hat{a}^\dagger_{M_k}(V_{u,v})|\Omega\rangle||C_1(M_1)|\cdots|C_k(M_k)|.$$

Clearly the matrix element is zero unless $M_j \subseteq \{u, v\}$ for all j. Expanding the commutators one gets 2^k terms, but only the term in which all creation operators stand on the right of $V_{u,v}$ gives a non-zero contribution. Taking into account that

$$\sum_{M_j\,:\,M_j\subseteq\{u,v\}} |C_j(M_j)| \leq 2\|C_j\|_1,$$

one arrives at

$$|\langle\Omega|\hat{C}_1\cdots\hat{C}_k(V_{u,v})|\Omega\rangle| \leq 2^k J \prod_{j=1}^{k} \|C_j\|_1 \leq 2^4 J \prod_{j=1}^{k} \|C_j\|_1,$$

where we have applied Lemma 2 to argue that $k \leq 4$.

References

[1] J. Kempe, A. Kitaev, and O. Regev, *"The Complexity of the Local Hamiltonian Problem"*, SIAM Journal of Computing, Vol. 35(5), p. 1070-1097 (2006).

[2] R. Oliveira and B. M. Terhal, *"The complexity of quantum spin systems on a two-dimensional square lattice"*, arXiv:quant-ph/0504050.

[3] T. Kato, *"Perturbation Theory for Linear Operators"*, Springer-Verlag New York (1966).

[4] D. Yarotsky, *"Perturbations of ground states in weakly interacting quantum spin systems"*, J. Math.Phys. **45**, N. 6, p. 2134 (2004).

[5] T. Osborne, *"Simulating adiabatic evolution of gapped spin systems"*, Phys. Rev. A **75**, 032321 (2007).

[6] A. Abrikosov, L. Gorkov, and I. Dzyaloshinski, *"Methods of Quantum Field Theory in Statistical Physics"*, Dover Publications Inc., New York (1975).

[7] I. Lindgren, *"The Rayleigh-Schrödinger perturbation and the linked-diagram theorem for a multi-configurational model space"*, J. Phys. B, Vol. 7, No. 18, p. 2441 (1974).

[8] J. Kirkwood and L. Thomas, "Expansions and Phase Transitions for the Ground State of Quantum Ising Lattice Systems", Commun. Math. Phys. **88**, pp. 569-580 (1983).

[9] N. Datta and T. Kennedy, *"Expansions for one quasiparticle states in spin 1/2 systems"*, J. Stat. Phys. **108**, p. 373 (2002).

[10] S. Lang, *"Complex Analysis"*, Graduate Texts in Mathematics 103, Springer-Verlag New York (1985).

[11] P. W. Anderson, *"Infrared Catastrophe in Fermi Gases with Local Scattering Potentials"*, Phys. Rev. Lett. **18**, p. 1049 (1967).

[12] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, *"Criticality, the area law, and the computational power of PEPS"*, Phys. Rev. Lett. **96**, 220601 (2006).

[13] F. Verstraete, *private communication* (2007).

[14] F. Coester, *"Bound states of a many-particle system"*, Nucl. Phys. **7**, p. 421 (1958).

[15] T. Crawford and H. Schaefer, *"An Introduction to Coupled Cluster Theory for Computational Chemists"*, Reviews in Computational Chemistry, Vol. 14 (1999).

[16] D.J.J. Farnell and R.F. Bishop, *"The Coupled Cluster Method Applied to the XXZ Model on the Square Lattice"*, arXiv:cond-mat/0606060.

[17] M. Hastings and T. Koma, *"Spectral Gap and Exponential Decay of Correlations"*, Commun. Math. Phys. **265**, p. 781 (2006).

[18] S. Bravyi, M. Hastings, and F. Verstraete, *"Lieb-Robinson bounds and the generation of correlations and topological quantum order"*, Phys. Rev. Lett. **97**, 050401 (2006).

[19] J. I. Latorre, E. Rico, and G. Vidal, *"Ground state entanglement in quantum spin chains"*, Quant. Inf. Comput. **4**, p. 48 (2004).

[20] R. Bhatia, *"Matrix Analysis"*, Graduate Texts in Mathematics 169, Springer-Verlag New York (1997).

[21] P. Aliferis, D. Gottesman, and J. Preskill, *"Accuracy threshold for postselected quantum computation"*, arXiv:quant-ph/0703264.