The Thouless theorem for matrix product states and subsequent post-density matrix renormalization group methods

Sebastian Wouters,[1,][*] Naoki Nakatani,[2] Dimitri Van Neck,[1] and Garnet Kin-Lic Chan[2]

[1]*Center for Molecular Modeling, Ghent University, Technologiepark 903, 9052 Zwijnaarde, Belgium*
[2]*Department of Chemistry, Princeton University,*
Frick Chemistry Laboratory, Princeton, NJ 08544, USA

The similarities between Hartree-Fock (HF) theory and the density-matrix renormalization group (DMRG) are explored. Both methods can be formulated as the variational optimization of a wavefunction ansatz. Linearization of the time-dependent variational principle near a variational minimum allows to derive the random phase approximation (RPA). We show that the non-redundant parametrization of the matrix product state (MPS) tangent space [J. Haegeman *et al.*, Phys. Rev. Lett. **107**, 070601 (2011)] leads to the Thouless theorem for MPS, i.e. an explicit non-redundant parametrization of the entire MPS manifold, starting from a specific MPS reference. Excitation operators are identified, which extends the analogy between HF and DMRG to the Tamm-Dancoff approximation (TDA), the configuration interaction (CI) expansion, and coupled cluster theory. For a small one-dimensional Hubbard chain, we use a CI-MPS ansatz with single and double excitations to improve on the ground state and to calculate low-lying excitation energies. For a symmetry-broken ground state of this model, we show that RPA-MPS allows to retrieve the Goldstone mode. We also discuss calculations of the RPA-MPS correlation energy. With the long-range quantum chemical Pariser-Parr-Pople Hamiltonian, low-lying TDA-MPS and RPA-MPS excitation energies for polyenes are obtained.

I. INTRODUCTION

The standard classification of quantum ground states dates back to Landau.[1] Mean-field theory is used to describe a state, and a phase transition is marked by the breaking of a symmetry. Particle-conserving mean-field theory for fermions is Hartree-Fock (HF) theory.[2] In HF theory, the exact ground state is approximated by a Slater determinant (SD),[3] and the energy of a Hamiltonian is minimized within this variational ansatz space. To obtain excited states or a more accurate description of the ground state, post-HF (post mean-field) methods[4] can be carried out such as the Tamm-Dancoff approximation (TDA),[5] the random-phase approximation (RPA),[6] Møller-Plesset perturbation theory,[7] the configuration interaction (CI) expansion,[3,8] and coupled cluster (CC) theory.[9]

Within the framework of second quantization,[10–12] the reference SD obtains a simple product form when the canonical HF orbitals are used to construct the Fock space. Occupied-virtual (OV) excitation operators allow to connect the reference SD to post-HF wavefunction ansatzes. The Thouless theorem gives a non-redundant parametrization to generate *all* possible SDs from any given SD reference, by means of its OV excitation operators.[13,14]

Recently, a new way to understand the qualitative structure of quantum many-body states has appeared, whereby the state is approximated by a tensor network, i.e. a contracted product of tensors where each tensor represents a local degree of freedom. These ansatzes are efficient representations of low-energy states because they capture the boundary law for the entanglement entropy. In one dimension, the tensor network is known as a matrix product state (MPS). The MPS is the wavefunction ansatz for the density-matrix renormalization group (DMRG) algorithm.[15]

DMRG can capture states beyond the realm of Landau (or mean-field) theory, i.e. states with topological order.[16] DMRG has also been shown to be a powerful method to treat the static correlation problem in electronic structure theory.[17–19] Static correlation arises when a state consists of several significant SD contributions, which HF theory is of course unable to deal with, because a single SD does not describe the qualitative structure of the targeted state. Post-HF methods, which start from a single SD reference, have difficulty building in large static correlation a posteriori. In these situations DMRG has provided a new ability to access the electronic structure. The analogue of static correlation for DMRG is a quantum critical system, which introduces corrections to the entanglement boundary law which cannot be captured by DMRG.

DMRG can be interpreted as a mean-field theory in the sites, which is analogous to HF, which is a mean-field theory in the particles.[20,21] Therefore, it is natural to search for extensions to DMRG that are analogous to post-HF methods: post-DMRG methods. One example is linear response theory. Time-dependent HF theory is obtained by using an SD ansatz in the time-dependent variational principle (TDVP).[13,22] Time-dependent DMRG (which stays within the MPS ansatz space) is similarly obtained by using an MPS ansatz in the TDVP.[21,23–25] RPA, or linear response theory for HF, is obtained by linearizing the time-dependent HF equations in the vicinity of a variational mimimum.[26] Equivalently, the RPA equations can be derived from an equation of motion (EOM) approach with excitation operators.[27] RPA yields a mean-field description of quasi-particle excitations. Linear response theory for DMRG was first derived by Dorando

et al.,[28] and was later recast as RPA for MPS.[21,25]

In this work, we construct a more complete analogue of the mean-field framework, which allows us to define a full set of post-DMRG methods. We give a non-redundant parametrization of the entire MPS manifold, starting from a specific MPS reference. This is the analogue of the Thouless theorem for HF. We identify the excitation operators of the Thouless theorem. These excitation operators allow for a complete rederivation of RPA for MPS by means of the EOM, in complete analogy with HF. All other results, such as an improvement of the ground state theory by the fluctuation-dissipation theorem, follow. With these excitation operators, we can define the analogues of other post-HF methods for MPS, such as CC and CI.

For a small one-dimensional Hubbard chain, we use a numerical CI-MPS ansatz with single and double excitations to improve on the ground state and to calculate low-lying excitation energies. For a symmetry-broken ground state of this model, we show that RPA-MPS allows to retrieve the Goldstone mode. We also discuss calculations of the RPA-MPS correlation energy. With the long-range quantum chemical Pariser-Parr-Pople (PPP) Hamiltonian, low-lying TDA-MPS and RPA-MPS excitation energies for polyenes are also obtained.

II. HF MEAN-FIELD THEORY

This section provides a brief introduction to the variational principles, HF mean-field theory, the Thouless theorem, and post-HF methods. It focusses on the topics for which a DMRG analogue will be constructed in this paper. For readers familiar with HF, this section can be a good guideline to understand our post-DMRG discussion.

A. Variational principles

Because the Hilbert space increases exponentially with system size, a variational wave-function ansatz $|\Phi(\mathbf{z})\rangle$ with parametrization $\mathbf{z}$ is often used to make calculations feasible. In order to minimize the energy functional

$$E(\mathbf{z}, \overline{\mathbf{z}}) = \frac{\langle \overline{\Phi} \mid \hat{H} \mid \Phi \rangle}{\langle \overline{\Phi} \mid \Phi \rangle} \quad (1)$$

to approximate ground states, the time-independent variational principle (TIVP) $\frac{\delta \mathcal{L}}{\delta \overline{z}} = 0$ can be employed, where

the Lagrangian is[20]

$$\mathcal{L} = \langle \overline{\Phi} \mid \hat{H} \mid \Phi \rangle - \lambda \left(\langle \overline{\Phi} \mid \Phi \rangle - 1 \right). \quad (2)$$

The overline denotes complex conjugation. This yields the time-independent self-consistent field (SCF) equations. To approximate time evolution, the time-dependent variational principle (TDVP) $\frac{\delta S}{\delta \overline{z}} = 0$ can be employed, where the action is[13,22,24]

$$S = \int_{t_1}^{t_2} dt \left(\frac{i\hbar}{2} \langle \overline{\Phi} \mid \dot{\Phi} \rangle - \frac{i\hbar}{2} \langle \dot{\overline{\Phi}} \mid \Phi \rangle - \langle \overline{\Phi} \mid \hat{H} \mid \Phi \rangle \right). \quad (3)$$

The dot denotes time derivation. This yields the time-dependent SCF equations.

B. The Slater determinant

From a given single-particle basis, any other single-particle basis can be constructed by a unitary transformation: $\hat{a}_j^\dagger = \hat{b}_k^\dagger U^k_{\ j}$. Second quantization is used to denote the single-particle states,[10–12] and the summation convention is used for double indices. An N-particle SD is an anti-symmetrized product of N single-particle states (called occupied orbitals):[3]

$$|\Psi\rangle = \hat{a}_1^\dagger \hat{a}_2^\dagger ... \hat{a}_N^\dagger |-\rangle \quad (4)$$

The variational freedom is a unitary transformation from the given single-particle basis of L orbitals to another basis where the first N orbitals are used to construct the SD. There is gauge freedom in the ansatz, as any unitary transformation that does not mix the N occupied orbitals with the $L-N$ virtual orbitals, does not change the wave-function (except for a global phase). An SD is therefore described by the Grassmann manifold $\mathsf{U}_L/(\mathsf{U}_N \times \mathsf{U}_{L-N})$, with U_k the unitary group of $k \times k$ unitary matrices. This manifold has dimension $2N(L-N)$, and can be parametrized by $N(L-N)$ complex numbers.[13] This will henceforth be called a complex dimension $N(L-N)$.

C. The Fock equations

If the particles of a system interact pairwise, the Hamiltonian can always be written in second quantization as

$$\hat{H} = \hat{b}_i^\dagger \mathsf{T}^i_{\ j} \hat{b}^j + \frac{1}{2} \hat{b}_i^\dagger \hat{b}_j^\dagger \mathsf{V}^{ij}_{\ kl} \hat{b}^l \hat{b}^k. \quad (5)$$

The TIVP can be expressed in terms of the unitary transformation generating the occupied orbitals:

$$\mathcal{L} = U^{\dagger \alpha}_{\ i} \mathsf{T}^i_{\ j} U^j_{\ \alpha} + \frac{1}{2} U^{\dagger \alpha}_{\ i} U^{\dagger \beta}_{\ j} \mathsf{V}^{ij}_{\ kl} U^l_{\ \beta} U^k_{\ \alpha} - \frac{1}{2} U^{\dagger \alpha}_{\ i} U^{\dagger \beta}_{\ j} \mathsf{V}^{ij}_{\ kl} U^l_{\ \alpha} U^k_{\ \beta} - \lambda^\beta_{\ \alpha} \left(U^{\dagger \alpha}_{\ k} U^k_{\ \beta} - \delta^\alpha_{\ \beta} \right). \quad (6)$$

The Greek indices denote occupied orbitals, while the Latin indices denote all single-particle basis states. Varying with respect to $U^{\dagger m}{}_i$, leads to the Fock equations:[2]

$$\mathsf{F}^i{}_k U^k{}_p = \left(\mathsf{T}^i{}_k + U^{\dagger\beta}{}_j \mathsf{V}^{ij}{}_{kl} U^l{}_\beta - U^{\dagger\beta}{}_j \mathsf{V}^{ij}{}_{lk} U^l{}_\beta \right) U^k{}_p = U^i{}_q \lambda^q{}_p. \tag{7}$$

The gauge can be partially fixed by requiring that the Lagrangian multiplier matrix λ to enforce orthonormal orbitals becomes diagonal, and that the diagonal elements are sorted in ascending order. These diagonal elements are then interpreted as the single-particle energy levels.[4,13] The remaining gauge freedom is then $\mathsf{U}_1^{\otimes L}$, i.e. the phase of each HF orbital. The lowest N single-particle states are used to construct the SD.

The Fock equations are orbital-based mean-field equations. There is self-consistency because the Fock operator in Eq. (7) which determines the orbitals, also depends on the orbitals.

D. The Thouless theorem

The Thouless theorem for HF[14] and its unitary counterpart[13] state that any N-electron SD can be globally parametrized as resp.

$$|\Psi\rangle \propto \exp\left(X^{vo} \hat{B}^\dagger_{vo} \right) |\Psi^0\rangle \tag{8}$$

$$|\Psi\rangle = \exp\left(X^{vo} \hat{B}^\dagger_{vo} - \overline{X}_{vo} \hat{B}^{vo} \right) |\Psi^0\rangle \tag{9}$$

with $|\Psi^0\rangle$ a random SD, with N occupied (o) and $L-N$ virtual (v) orbitals. $\hat{B}^\dagger_{vo}$ is a shorthand for $\hat{a}^\dagger_v \hat{a}^o$. Note that the summation convention was used. The equality in Eq. (9) holds because the exponential of an anti-hermitian operator is unitary, and hence does not change the norm.

This parametrization is of complex dimension $N(L-N)$, and is thus non-redundant. All parameters X are needed to parametrize the neighbourhood of $|\Psi^0\rangle$. For all possible combinations $|\Psi^0\rangle$ and $|\Psi\rangle$, a solution X can always be found. The theorem does not state that this X is unique. In fact, X is not unique, see e.g. the discussion in the appendix of Rowe *et al.*[13] The reader can think about the Lie group of $\mathsf{O}(3)$, where several combinations of successive rotations along different axes can generate the same global rotation. Instead of working with the redundant parameters U, we can equivalently work with the non-redundant parameters X.

The nth order variation of a wave-function defines its nth order tangent space. The (first order) tangent space of this non-redundant parametrization consists of the single OV excitations $\hat{B}^\dagger_{vo} |\Psi^0\rangle$.

E. Time evolution

The TDVP leads to the time-dependent SCF equations:[13,22,29]

$$i\hbar \dot{U}(t)^i{}_p = F(t)^i{}_k U(t)^k{}_p \tag{10}$$

The Fock operator dictates how orbitals are rotated into each other over time. Rotations within the space of occupied orbitals or within the space of virtual orbitals, do not change the SD wavefunction as it represents a Grassmann manifold. Only the rotation of occupied and virtual orbitals into each other has physical meaning.

To obtain the rate of OV rotation determined by Eq. (10) in the point $|\Psi^0\rangle$, the Thouless parametrization of a general SD can be used in the TDVP:

$$i\hbar \dot{X}^{vo}(\mathbf{X} = \mathbf{0}) = \langle \overline{\Psi^0} | \hat{B}^{vo} \hat{H} | \Psi^0 \rangle. \tag{11}$$

The parameters X^{vo} are flattened to a column $\mathbf{X}$. The same equation is obtained by inserting Eq. (9) in the time-dependent Schrödinger equation, and by projecting this equation onto $\hat{B}^\dagger_{vo} |\Psi\rangle = \hat{B}^\dagger_{vo} |\Psi(\mathbf{X}, \overline{\mathbf{X}})\rangle$. The time evolution and its projection are resp. given by

$$i\hbar \left(\dot{X}^{wp} \frac{\partial}{\partial X^{wp}} + \dot{\overline{X}}_{wp} \frac{\partial}{\partial \overline{X}_{wp}} \right) |\Psi\rangle = \left(\hat{H} - E_{\text{HF}} \right) |\Psi\rangle \tag{12}$$

$$i\hbar \langle \overline{\Psi}| \hat{B}^{vo} \left(\dot{X}^{wp} \frac{\partial}{\partial X^{wp}} + \dot{\overline{X}}_{wp} \frac{\partial}{\partial \overline{X}_{wp}} \right) |\Psi\rangle$$
$$= \langle \overline{\Psi}| \hat{B}^{vo} \left(\hat{H} - E_{\text{HF}} \right) |\Psi\rangle \tag{13}$$

Evaluation for $\mathbf{X} = \mathbf{0}$ yields Eq. (11).

F. RPA

Linearization of the TDVP near a variational minimum leads to RPA.[13,26,29] Take the variational minimum as the reference $|\Psi^0\rangle$. Expand Eq. (13) up to first order around $\mathbf{X} = \mathbf{0}$. The zeroth order terms vanish because the expansion point is a variational minimum: $\langle \overline{\Psi^0} | \hat{B}^{vo} \hat{H} | \Psi^0 \rangle = 0$. This is Brillouin's theorem.[30] The linearized equations are:

$$i\hbar \dot{X}^{vo} = \langle \overline{\Psi^0} | \hat{B}^{wp} \hat{B}^{vo} \left(\hat{H} - E_{\text{HF}} \right) | \Psi^0 \rangle \overline{X}_{wp}$$
$$+ \langle \overline{\Psi^0} | \hat{B}^{vo} \left(\hat{H} - E_{\text{HF}} \right) \hat{B}^\dagger_{wp} | \Psi^0 \rangle X^{wp} \tag{14}$$

$$= - \langle \overline{\Psi^0} | \left[\hat{B}^{vo}, \left[\hat{H}, \hat{B}^{wp} \right] \right] | \Psi^0 \rangle \overline{X}_{wp}$$
$$+ \langle \overline{\Psi^0} | \left[\hat{B}^{vo}, \left[\hat{H}, \hat{B}^\dagger_{wp} \right] \right] | \Psi^0 \rangle X^{wp} \tag{15}$$

Assume a harmonic motion of the form $\mathbf{X} = \mathbf{Y}e^{-i\omega t} + \overline{\mathbf{Z}}e^{i\omega t}$. This leads to the RPA equations:

$$\hbar\omega \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} = \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} \qquad (16)$$

with $A_{vo;wp} = \langle \overline{\Psi^0} \mid [\hat{B}^{vo}, [\hat{H}, \hat{B}_{wp}^\dagger]] \mid \Psi^0\rangle$ and $B_{vo;wp} = -\langle \overline{\Psi^0} \mid [\hat{B}^{vo}, [\hat{H}, \hat{B}^{wp}]] \mid \Psi^0\rangle$. Note that if $(\omega, \mathbf{Y}, \mathbf{Z})$ is a solution, $(-\omega, \overline{\mathbf{Z}}, \overline{\mathbf{Y}})$ is a solution too.

Consider the energy functional

$$E(\mathbf{X}, \overline{\mathbf{X}}) = \langle \overline{\Psi(\mathbf{X}, \overline{\mathbf{X}})} \mid \hat{H} \mid \Psi(\mathbf{X}, \overline{\mathbf{X}})\rangle \qquad (17)$$

and its expansion up to second order in $\mathbf{X}$:

$$E^{(2)}(\mathbf{X}, \overline{\mathbf{X}}) - E_{\text{HF}} = \frac{1}{2} \begin{pmatrix} \mathbf{X} \\ \overline{\mathbf{X}} \end{pmatrix}^\dagger \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ \overline{\mathbf{X}} \end{pmatrix}. \qquad (18)$$

The RPA method searches for the harmonic modes of this potential near its minimum, akin to normal mode analysis in analytical mechanics.

In linear response theory, the RPA frequencies occur as poles in the response function. Because the exact response function for the exact ground state has the excitation energies of the Hamiltonian as poles, the RPA frequencies are interpreted as approximate excitation energies.[29] A second argument to interpret the RPA frequencies as excitation energies is given by the alternative derivation of RPA by means of the EOM.[27,29] Assume we know the exact ground state $|0\rangle$, and the exact excitation operators which connect the ground state to the excited states $\hat{Q}_n^\dagger = |n\rangle \langle 0|$. The operator $\hat{Q}_n$ then destroys the ground state. With $\hbar\omega_n = E_n - E_0$, the excitation energy of the excited state $|n\rangle$, it is easy to derive the EOM:

$$\langle \overline{0} \mid [\delta\hat{Q}, [\hat{H}, \hat{Q}_n^\dagger]] \mid 0\rangle = \hbar\omega_n \langle \overline{0} \mid [\delta\hat{Q}, \hat{Q}_n^\dagger] \mid 0\rangle. \qquad (19)$$

For RPA, two assumptions are made: the excitation operators are approximated by $\hat{Q}_n^\dagger = Y_n^{vo}\hat{B}_{vo}^\dagger - Z_n^{vo}\hat{B}^{vo}$, and the expectation values of the commutators are calculated with the HF reference wavefunction. The latter approximation is called the quasi-boson approximation. The RPA Eq. (16) is then retrieved.

As an alternative, an exact bosonic algebra can be set up:

$$[\hat{B}^{vo}, \hat{B}_{wp}^\dagger] = \delta_w^v \delta_p^o \qquad (20)$$

$$[\hat{B}^{vo}, \hat{B}^{wp}] = 0 \qquad (21)$$

by adding higher order terms to $\hat{B}_{vo}^\dagger = \hat{a}_v^\dagger \hat{a}_o + \mathcal{O}(\hat{a}^4)$. The Hamiltonian can then be written in terms of these new operators:

$$\hat{H}_B = E_{\text{HF}} - \frac{\text{tr}A}{2} + \frac{1}{2} \left(\hat{\mathbf{B}}^\dagger \hat{\mathbf{B}}\right) \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{pmatrix} \hat{\mathbf{B}} \\ \hat{\mathbf{B}}^\dagger \end{pmatrix} + \mathcal{O}(\hat{B}^3). \qquad (22)$$

RPA coincides with neglecting all terms of $\mathcal{O}(\hat{B}^3)$ in the bosonic expansion. This leads to the RPA correlation energy and wavefunction:[14,29]

$$E_{\text{cRPA}} = -\frac{\text{tr}A}{2} + \sum_{\omega_n > 0} \frac{\hbar\omega_n}{2} = -\sum_n \hbar\omega_n \sum_{vo} \mid Z_n^{vo} \mid^2 \quad (23)$$

$$|\text{RPA}\rangle \propto e^{\frac{1}{2}\left(\overline{ZY^{-1}}\right)^{vo;wp}\hat{B}_{vo}^\dagger\hat{B}_{wp}^\dagger} |\Phi^0\rangle. \qquad (24)$$

The RPA correlation energy has contributions from the zero point energy of the harmonic oscillators with frequency ω_n. The RPA wavefunction vanishes by the action of deexcitation operators: $\hat{Q}_n |\text{RPA}\rangle = 0$.

If the Hamiltonian has a continuous symmetry, and the exact ground state is degenerate due to this symmetry, a ground state calculation typically breaks this symmetry. Think for example about a spin-$\frac{1}{2}$ ground state. A calculation will lead to one possibility: $\alpha |s^z = \frac{1}{2}\rangle + \beta |s^z = -\frac{1}{2}\rangle$. A gapless bosonic degree of freedom remains, which corresponds to rotating within the spin-$\frac{1}{2}$ multiplet, called a Goldstone boson.[31] An interesting feature of RPA is its ability to retrieve Goldstone modes. The excitation energy of a Goldstone mode is of course zero, and the mode is its own dual solution $(\omega = 0, \mathbf{Y}, \mathbf{Z}) = (\omega = 0, \overline{\mathbf{Z}}, \overline{\mathbf{Y}})$.[12,14] This implies that

$$\sum_{vo} \left(\overline{Y}_{vo}Y^{vo} - \overline{Z}_{vo}Z^{vo}\right) = 0. \qquad (25)$$

G. Post-HF methods

With the excitation operators $\hat{B}_{vo}^\dagger = \hat{a}_v^\dagger \hat{a}^o$, a set of orthonormal vectors can be generated: $|\text{HF}\rangle$, $\hat{B}_{vo}^\dagger |\text{HF}\rangle$, $\hat{B}_{vo}^\dagger\hat{B}_{wp}^\dagger |\text{HF}\rangle$... They correspond to the zeroth, first and second order tangent space of the Thouless parametrization of a general SD. With the CI method, eigenstates of the Hamiltonian are approximated by working in an incomplete basis of such vectors.[3,8] Consider for example the second order expansion CISD, or CI with single and double excitations:

$$|\text{CISD}\rangle \propto \left(x + y^{vo}\hat{B}_{vo}^\dagger + \frac{1}{2}z^{vo;wp}\hat{B}_{vo}^\dagger\hat{B}_{wp}^\dagger\right) |\text{HF}\rangle \quad (26)$$

With CIS, or CI with only single excitations, the lowest energy state is again $|\text{HF}\rangle$ due to Brillouin's theorem, and the eigenstates approximated in the basis $\hat{B}_{vo}^\dagger |\text{HF}\rangle$ are therefore excited states. Note that this corresponds to diagonalizing the A-matrix of RPA in Eq. (16). Historically, this method is known as TDA.[5]

The RPA wavefunction in Eq. (24) suggests a CC ansatz.[9] Consider for example CCSD, or CC with single and double excitations:

$$|\text{CCSD}\rangle \propto e^{\left(y^{vo}\hat{B}_{vo}^\dagger + \frac{1}{2}z^{vo;wp}\hat{B}_{vo}^\dagger\hat{B}_{wp}^\dagger\right)} |\text{HF}\rangle \qquad (27)$$

An important property of ansatz wave-functions is their size consistency, i.e. the property that for two non-interacting subsystems, the compound wave-function is

multiplicatively separable and the total energy additively separable. CISD is not size consistent if there are more than two electrons in the compound system, whereas CCSD is always size consistent because of the exponential ansatz.[4]

III.   THE MATRIX PRODUCT STATE

A.   The ansatz

Consider the many-body Hilbert space $|n_1 n_2 ... n_L\rangle$, formed by taking the direct product of L local Hilbert spaces $|n_i\rangle$. The local degrees of freedom can be e.g. the spin projections of spins on a lattice, or the occupancies of orbitals. In the latter case, the states $|n_1 n_2 ... n_L\rangle$ form the Fock space.[11] An MPS can be seen as a linear combination of these vectors, where the coefficient of each vector is a product of matrices:

$$|\Phi\rangle = \sum_{\{n_i\}} A[1]^{n_1} A[2]^{n_2} ... A[L]^{n_L} |n_1 n_2 ... n_L\rangle. \quad (28)$$

We assume an MPS with open boundary conditions, i.e. the first matrix has row dimension 1 and the last matrix has column dimension 1. The bond dimension (virtual dimension) D_i of an MPS at boundary i is the column dimension of the matrices at site i and the row dimension of the matrices at site $i + 1$. With our assumption, $D_0 = D_L = 1$. The total number of complex parameters in this ansatz is $\dim\mathbb{A} = \sum_{i=1}^{L} D_{i-1} d D_i$, with d the size of the local Hilbert space $|n_i\rangle$. Just as in the SD, there is gauge freedom in the ansatz. Right-multiplying the d site-matrices on site i with the non-singular matrix G ($\tilde{A}[i]^{n_i} = A[i]^{n_i} G$), and simultaneously left-multiplying the d site-matrices on site $i + 1$ with its inverse G^{-1} ($\tilde{A}[i+1]^{n_{i+1}} = G^{-1} A[i+1]^{n_{i+1}}$), does not change the wave-function ($\tilde{A}[i]^{n_i} \tilde{A}[i+1]^{n_{i+1}} = A[i]^{n_i} A[i+1]^{n_{i+1}}$). A global scalar multiplication does not change the wave-function either. The MPS manifold, i.e. the quotient space of the general parametrization (complex dimension $\dim\mathbb{A}$) and all gauge freedom (complex dimension $\sum_{i=1}^{L} D_i^2$), has complex dimension

$\dim\mathbb{T} = \sum_{i=1}^{L} (d D_{i-1} - D_i) D_i$.[23,32,33]

B.   The SD as low bond dimension limit

An interesting connection to HF can be made by considering an MPS where the L orbitals are the HF orbitals. As each orbital occupation number is definite in an SD, an MPS with bond dimension 1 suffices to represent it. Conversely, if an MPS has bond dimension 1 and represents a state with definite particle number, each orbital has a definite occupation number. If this is not the case, two or more orbitals must be entangled (there is static correlation between them), and the bond dimension has to be larger than 1 to represent this. An MPS with bond dimension 1 and definite particle number can hence always be represented by an SD. An SD is the low bond dimension limit of an MPS, while a general full CI (FCI) solution requires an exponentially large bond dimension to be represented by an MPS.[34]

The SD ansatz provides a single variational approximation to the ground state, which unfortunately fails to represent static correlation. On the contrary, the MPS ansatz allows to systematically improve the approximation to the ground state by increasing the bond dimension, up to the point where all static correlation is resolved.[17–19]

IV.   THE DMRG EQUATIONS

The TIVP leads to the DMRG equations.[20] The canonical DMRG equations for site i are retrieved when additional constraints are added to the Lagrangian to enforce that the site-matrices to the left of site i are left-normalized:

$$\forall j < i : \sum_{n_j} (A^{n_j}[j])^\dagger A^{n_j}[j] = I_{D_j} \quad (29)$$

and that the site-matrices to the right of site i are right-normalized:

$$\forall j > i : \sum_{n_j} A^{n_j}[j] (A^{n_j}[j])^\dagger = I_{D_{j-1}}. \quad (30)$$

With $(A^{n_j}[j])_{\alpha\beta} = A[j]^{n_j \alpha\beta}$, the Lagrangian becomes

$$\mathcal{L} = \langle \overline{\Phi} | \hat{H} | \Phi \rangle - \lambda \left(\overline{A[i]_{n_i \alpha\beta}} A[i]^{n_i \alpha\beta} - 1 \right) - \sum_{j<i} \lambda[j]_\gamma^{\ \beta} \left(\overline{A[j]_{n_j \alpha\beta}} A[j]^{n_j \alpha\gamma} - \delta_\beta^{\ \gamma} \right) - \sum_{j>i} \lambda[j]_\gamma^{\ \alpha} \left(\overline{A[j]_{n_j \alpha\beta}} A[j]^{n_j \gamma\beta} - \delta_\alpha^{\ \gamma} \right). \quad (31)$$

Varying with respect to $\overline{A[i]_{n_i \alpha\beta}}$ gives the canonical one-site DMRG equations:

$$H_{\text{eff}}[i]^{n_i \alpha\beta}_{\ \ \tilde{n}_i \tilde{\alpha} \tilde{\beta}} A[i]^{\tilde{n}_i \tilde{\alpha} \tilde{\beta}} = \lambda A[i]^{n_i \alpha\beta} \quad (32)$$

in terms of the effective Hamiltonian.[20] By bringing the MPS into canonical forms, of which the left- and right-normalization conditions above are examples, the gauge freedom can be (partially) removed. For the left- and

right-normalization conditions, the remaining gauge freedom is a unitary rotation (G unitary). All gauge freedom can be removed by bringing the MPS into Vidal's canonical form.[35]

The DMRG equations are site-based mean-field equations. There is self-consistency because the effective Hamiltonian in Eq. (32) which determines the site-matrices of a particular site, depends on the site-matrices of the other sites.[20,21,34] In DMRG, the effective Hamiltonian hence plays the role of Fock operator.[20] Since both of them act locally (resp. on one site and one orbital), it might be worthwhile to explore Rayleigh-Schrödinger perturbation theory analogues for DMRG in the future, such as Møller-Plesset perturbation theory.[4,7,20]

Note that in practice the two-site DMRG algorithm is used to optimize an MPS. The two-site algorithm is more robust against local minima, and when symmetry is imposed it provides a natural way to distribute the bond dimension D over the symmetry sectors. After the two-site algorithm has converged, a few one-site DMRG sweeps allow to make the MPS fully self-consistent. This can be compared to HF, where the optimal SD is found by gradient methods[29] or by direct inversion of iterative subspaces[36] for stability reasons. The DMRG and HF solutions satisfy resp. Eqs. (32) and (7), irrespective of the optimization scheme.

V. THE MPS TANGENT SPACE

A. A redundant parametrization

Flatten the site-matrices $A[i]^{n_i}$ to a column $\mathbf{A}$ with entries $(A[i]^{n_i})_{\alpha,\beta} = A^{in_i\alpha\beta} = A^\mu$, and consider a small variation $A^\mu = A_0^\mu + B^\mu$. The wave-function can then be expanded as

$$|\Phi\rangle = |\Phi^0\rangle + B^\mu |\Phi_\mu^0\rangle + \frac{1}{2} B^\mu B^\nu |\Phi_{\mu\nu}^0\rangle + ... \quad (33)$$

with first order tangent space $|\Phi_\mu^0\rangle = \partial_\mu |\Phi^0\rangle = \frac{\partial |\Phi^0\rangle}{\partial A^\mu}$ and second order tangent space $|\Phi_{\mu\nu}^0\rangle = \partial_\mu \partial_\nu |\Phi^0\rangle$. Note that the summation convention was used. Each order of MPS tangent space contains all lower orders: e.g. $A_0^\mu |\Phi_\mu^0\rangle = L |\Phi^0\rangle$ and $A_0^\mu B^\nu |\Phi_{\mu\nu}^0\rangle = (L-1) B^\mu |\Phi_\mu^0\rangle$.[21,23,32]

The tangent vectors $|\Phi_\mu^0\rangle$ are redundant, as the MPS manifold has dimension dim$\mathbb{T}$, and there are dim$\mathbb{A}$ such vectors. The metric, or overlap matrix $S_{\mu\nu} = \langle \overline{\Phi_\mu^0} \mid \Phi_\nu^0 \rangle$, is therefore not invertible. In Sec. V C, dim$\mathbb{T}$ explicit linear combinations of the vectors $|\Phi_\mu^0\rangle$ are given, so that the overlap in this new basis is the unit matrix, and $|\Phi^0\rangle$ is orthogonal to this new basis. Remember that variations in the direction of $|\Phi^0\rangle$ only cause norm or phase changes of the ansatz, but do not change the physical state. This new basis is then a non-redundant parametrization of the MPS tangent space.

B. Hamiltonian sparsity

The Hamiltonian in Eq. (5) is sparse, as it consists of a sum of one- and two-particle interactions. When it acts on a certain SD, the result lies in the space spanned by the given SD and its single and double OV excitations. This is immediataly clear by changing the single particle basis in Eq. (5) from $\hat{b}_k^\dagger$ to the SD orbitals $\hat{a}_j^\dagger$.

A typical lattice Hamiltonian can be considered sparse too, as it consists of a sum of one- and two-site operators. It is sparse in site-space instead of particle-space. Let us focus on the one-dimensional Hubbard model:[37]

$$\hat{H} = - \sum_{\sigma,i=1}^{L-1} \left(\hat{a}_{i\sigma}^\dagger \hat{a}_{i+1\sigma} + \hat{a}_{i+1\sigma}^\dagger \hat{a}_{i\sigma} \right) + U \sum_{i=1}^{L} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}. \quad (34)$$

Consider its action on an MPS. Let μ_i be a shorthand for (n_i, α, β), or μ restricted to site i. The Hamiltonian connects the MPS to a part of its double tangent space:

$$\hat{H} |\Phi^0\rangle \propto C^{\mu_i \nu_{i+1}} |\Phi_{\mu_i \nu_{i+1}}^0\rangle. \quad (35)$$

It might hence be worthwhile to construct the site-space analogue of the particle Fock space.[10,11] A new second quantization should be constructed, based on the MPS reference instead of the HF orbitals.

C. A non-redundant parametrization

A non-redundant parametrization of the MPS tangent space was first presented by Dorando *et al.*[28] in DMRG projector terminology. Haegeman *et al.*[23] provided a construction in the language of the MPS wave-function and the corresponding manifold. To present the relationship between the two, here we describe the tangent space construction in projector terms, but by using the explicit MPS representation of the projectors.

Consider an MPS where all left renormalized basis states at boundary $i-1$:

$$|L_\alpha^{i-1}\rangle = \sum_{\{n_j : j < i\}} [A^{n_1}[1]..A^{n_{i-1}}[i-1]]_\alpha |n_1..n_{i-1}\rangle \quad (36)$$

are orthonormal and all right renormalized basis states at boundary i:

$$|R_\beta^i\rangle = \sum_{\{n_j : j > i\}} [A^{n_{i+1}}[i+1]..A^{n_L}[L]]_\beta |n_{i+1}..n_L\rangle \quad (37)$$

are orthonormal. In the DMRG algorithm, a renormalization transformation is constructed to reduce the direct product of $|L_\alpha^{i-1}\rangle$ (size D_{i-1}) and $|n_i\rangle$ (size d) to a new left renormalized basis at boundary i (size $D_i \leq dD_{i-1}$). This renormalization transformation is a projection, represented by the site-matrices of site i:

$$\sum_{\alpha n_i} (A[i]^{n_i})_{\alpha,\beta} |L_\alpha^{i-1}\rangle |n_i\rangle. \quad (38)$$

The projection onto the $dD_{i-1} - D_i$ discarded states from the direct product space, defines the non-redundant tangent space. We now explain the explicit construction of the non-redundant tangent space as provided by Dorando *et al.* [28] in MPS terminology. Consider the QR-decomposition of the projector:

$$(A[i]^{n_i})_{\alpha,\beta} = A[i]_{(\alpha n_i),\beta} = \sum_{\gamma} Q[i]_{(\alpha n_i),\gamma} R_{\gamma,\beta}. \quad (39)$$

$Q[i]$ contains D_i orthonormal columns of size dD_{i-1}. Its left nullspace is spanned by $dD_{i-1} - D_i$ vectors. This allows to construct the $dD_{i-1} \times (dD_{i-1} - D_i)$ matrix $\tilde{Q}[i]$, so that $\left[Q[i] \tilde{Q}[i] \right]$ is unitary. A part of the non-redundant tangent space can then be parametrized by the matrix $x[i]$ with dimensions $(dD_{i-1} - D_i) \times D_i$:

$$\sum_{\alpha\beta n_i} \left(\tilde{Q}[i]^{n_i} x[i] \right)_{\alpha,\beta} |L_\alpha^{i-1}\rangle |n_i\rangle |R_\beta^i\rangle. \quad (40)$$

If the renormalized basis states $|L_\alpha^{i-1}\rangle$, $|R_\beta^i\rangle$ are not orthonormal, their overlap has to be taken into account. It was Haegeman *et al.* [23] who first presented the parametrization in that case:

$$\sum_{\alpha\beta n_i} \left(l[i-1]^{-\frac{1}{2}} \tilde{Q}[i]^{n_i} x[i] r[i]^{-\frac{1}{2}} \right)_{\alpha,\beta} |L_\alpha^{i-1}\rangle |n_i\rangle |R_\beta^i\rangle$$
$$(41)$$

with $l[i-1]$ the density matrix of the left renormalized states $|L_\alpha^{i-1}\rangle$ and $r[i]$ the density matrix of the right renormalized states $|R_\beta^i\rangle$. The QR-decomposition of Eq. (39) is now performed on $l[i-1]^{\frac{1}{2}} A[i]^{n_i}$ instead of on $A[i]^{n_i}$. The complete non-redundant tangent space is formed by doing this construction for the projector on each site. Combine all matrices $x[i]$ to a column $\mathbf{x}$ of length $\dim\mathbb{T}$. By writing the construction in Eq. (41) as $B^\mu(\mathbf{x}) |\Phi_\mu^0\rangle$, with

$$B^{n_i}(\mathbf{x})[i] = l[i-1]^{-\frac{1}{2}} \tilde{Q}^{n_i}[i] x[i] r[i]^{-\frac{1}{2}}, \quad (42)$$

one possibility for a non-redundant tangent space basis of dimension $\dim\mathbb{T}$ is immediately obtained:

$$|\Phi_k^T\rangle = \frac{\partial}{\partial x^k} B^\mu(\mathbf{x}) |\Phi_\mu^0\rangle. \quad (43)$$

Note that this provides a construction of $|\Phi_k^T\rangle$ as a linear combination of $|\Phi_\mu^0\rangle$. Any tangent vector can be constructed by taking a complex linear combination of these $\dim\mathbb{T}$ vectors: $x^k |\Phi_k^T\rangle = B^\mu(\mathbf{x}) |\Phi_\mu^0\rangle$. Because of the construction of $\tilde{Q}^{n_i}[i]$, these vectors are orthogonal to $|\Phi^0\rangle$: $\langle \overline{\Phi^0} | \Phi_\mu^0 \rangle B^\mu(\mathbf{x}) = 0$. The metric of the parametrization in Eq. (42) is the unit matrix: $\overline{B^\mu(\mathbf{x})} S_{\mu\nu} B^\nu(\mathbf{y}) = \mathbf{x}^\dagger \mathbf{y}$.[23,32] Analogous results have been obtained in a different context.[33]

For an SD written as an MPS ($D = 1$ and $d = 2$), the non-redundant tangent space vectors correspond to the addition (removal) of an electron to (from) the system.

VI. THE THOULESS THEOREM FOR MPS

The operators $\hat{B}_{vo}^\dagger$ link an SD $|\Psi^0\rangle$ to its non-redundant tangent space $\hat{B}_{vo}^\dagger |\Psi^0\rangle$. Exponentiation of these operators led to the Thouless theorem. Here we present the MPS counterpart.

A. Proposal

For the sake of simplicity, we use a part of the gauge freedom to work with a left-canonical MPS. The left-normalization condition in Eq. (29) then holds for all sites. This implies $\forall i : l[i] = I_{D_i}$. Consider the following matrix notation for site-matrices: $C[i]$ has entries $(C[i])_{(\alpha n_i),\beta}$, i.e. the row-index of the matrices $C[i]$ contains the physical index n_i. The left-normalization condition then becomes

$$A[i]^\dagger A[i] = I_{D_i}. \quad (44)$$

Because of the construction of $\tilde{Q}[i]$, the site-matrices $B(\mathbf{x})[i]$ are left-orthogonal to the site-matrices $A[i]$:

$$B(\mathbf{x})[i]^\dagger A[i] = 0. \quad (45)$$

This allows to propose the MPS counterpart of the Thouless theorem:

$$A(\mathbf{x},\overline{\mathbf{x}})[i] = \exp \left(B(\mathbf{x})[i] A_0[i]^\dagger - A_0[i] B(\mathbf{x})[i]^\dagger \right) A_0[i]$$
$$(46)$$

where now $A_0[i]$ is used for $A[i]$ to clearly mark the difference with $A(\mathbf{x},\overline{\mathbf{x}})[i]$. The matrix in the exponential is anti-Hermitian, and the transformation in Eq. (46) is therefore unitary. As the $A_0[i]$ site-matrices were left-normalized, the $A(\mathbf{x},\overline{\mathbf{x}})[i]$ site-matrices are also left-normalized. An MPS built with the $A(\mathbf{x},\overline{\mathbf{x}})[i]$ site-matrices:

$$|\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle = \sum_{\{n_i\}} A(\mathbf{x},\overline{\mathbf{x}})[1]^{n_1} ... A(\mathbf{x},\overline{\mathbf{x}})[L]^{n_L} |n_1 n_2 ... n_L\rangle$$
$$(47)$$

is hence still left-canonical and therefore normalized. For $\mathbf{x} = \mathbf{0}$, $|\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle = |\Phi^0\rangle$. The tangent space of this MPS parametrization is familiar too:

$$\left. \frac{\partial}{\partial x^k} |\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle \right|_{\mathbf{x}=\mathbf{0}} = |\Phi_k^T\rangle \quad (48)$$

which can be easily checked by using Eqs. (44) and (45). $|\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle$ is therefore an explicit non-redundant parametrization of the MPS manifold in the neighbourhood of $|\Phi^0\rangle$.

B. Global validity

Here we show that Eq. (47) is a global parametrization of the MPS manifold, or that any MPS with bond

dimensions D_i can be generated from $|\Phi^0\rangle$ (which has the same bond dimensions). This implies that we can optimize over the parameters $\mathbf{x}$ instead of over $\mathbf{A}$ to find an energy minimum.

For a specific site index i, the parametrization $A(\mathbf{x}, \overline{\mathbf{x}})[i]$ of Eq. (46) is a Grassmann manifold with matrix dimensions $dD_{i-1} \times D_i$. Define $\mathbf{y}$ by $x[i] = y[i]r[i]^{\frac{1}{2}}$ to obtain

$$\tilde{A}(\mathbf{y}, \overline{\mathbf{y}}) = A(\mathbf{x}, \overline{\mathbf{x}})[i]$$
$$= \exp\left(\tilde{Q}[i]yQ[i]^\dagger - Q[i]y^\dagger \tilde{Q}[i]^\dagger\right)Q[i]. \quad (49)$$

Note the close analogy to Eq. (9). We show in the Appendix that Eq. (49) represents a Grassmann manifold. Note that we assume that the density matrix $r[i]$ is non-singular (i.e. $r[i]^{-\frac{1}{2}}$ exists) for the construction of the non-redundant tangent space in Eq. (42). For a left-canonical MPS, the Schmidt values are the positive square roots of the eigenvalues of $r[i]$. The condition of non-singular density matrices $r[i]$ is therefore equal to having all Schmidt values of $|\Phi^0\rangle$ non-zero. This is a condition for the global validity of Thouless's theorem for MPS.

Now give a normalized MPS $|\Phi^1\rangle$, with the only restriction that it has the same bond dimensions as $|\Phi^0\rangle$. We will prove by construction

$$\exists \mathbf{x} : \left|\langle\overline{\Phi^1} \mid \Phi(\mathbf{x}, \overline{\mathbf{x}})\rangle\right| = 1. \quad (50)$$

1. Set $i = 1$.

2. Use a part of the gauge freedom at boundary i to bring the site-matrices $A_1[i]$ of $|\Phi^1\rangle$ in left-normalized form: $A_1^L[i]$.

3. Find $x_1[i]$ so that the columns of $A(\mathbf{x_1}, \overline{\mathbf{x_1}})[i]$ and the columns of $A_1^L[i]$ span the same space, which is always possible because $A(\mathbf{x}, \overline{\mathbf{x}})[i]$ is a Grassmann manifold.

4. Use the remaining gauge freedom at boundary i, i.e. a unitary transformation U_{D_i}, to enforce $A(\mathbf{x_1}, \overline{\mathbf{x_1}})[i] = A_1^{\text{exact}}[i] = A_1^L[i]U_{D_i}$.

5. If $i < L$, set $i = i + 1$ and go to 2.

When the construction is finished, all parameters of $\mathbf{x_1}$ are assigned, and the gauge freedom in $|\Phi^1\rangle$ was used to write $|\Phi^1\rangle$ exactly as $|\Phi(\mathbf{x_1}, \overline{\mathbf{x_1}})\rangle$, i.e. $\forall i : A_1^{\text{exact}}[i] = A(\mathbf{x_1}, \overline{\mathbf{x_1}})[i]$. See Ref. 33 on the diffeomorphism between a finite chain MPS manifold and a product manifold of Grassmann manifolds. This concludes the proof that Eq. (47) can represent any MPS with the same bond dimensions, as long as $|\Phi^0\rangle$ does not have any vanishing Schmidt values. Note that the theorem guarantees a solution $\mathbf{x_1}$, but does not guarantee that this solution is unique, in analogy with the discussion in Sec. II D.

C. The double tangent space

To get a better understanding of the MPS double tangent space, consider the second order term of $A(\mathbf{x}, \overline{\mathbf{x}})[i]$:

$$A(\mathbf{x}, \overline{\mathbf{x}})[i] - A_0[i] - B(\mathbf{x})[i]$$
$$= -\frac{1}{2}A_0[i]B(\mathbf{x})[i]^\dagger B(\mathbf{x})[i] + \mathcal{O}(x^3)$$
$$= -\frac{1}{2}A_0[i]r[i]^{-\frac{1}{2}}x[i]^\dagger x[i]r[i]^{-\frac{1}{2}} + \mathcal{O}(x^3) \quad (51)$$

The expansion of $|\Phi(\mathbf{x}, \overline{\mathbf{x}})\rangle$ up to second order then consists of

1. The MPS reference $|\Phi(\mathbf{0}, \mathbf{0})\rangle = |\Phi^0\rangle$.

2. The tangent space $\frac{\partial}{\partial x^k}|\Phi(\mathbf{x}, \overline{\mathbf{x}})\rangle\big|_{\mathbf{x}=\mathbf{0}} = \frac{\partial}{\partial x^k}|\Phi(\mathbf{0}, \mathbf{0})\rangle = |\Phi_k^T\rangle$. Note that $\frac{\partial}{\partial \overline{x_k}}|\Phi(\mathbf{0}, \mathbf{0})\rangle = 0$. The tangent space consists of all possible connections between the unused basis states from $|L_\alpha^{i-1}\rangle \otimes |n_i\rangle$ and $|R_\beta^i\rangle$.

3. The non-local part of the double tangent space $\frac{\partial^2}{\partial x^k \partial x^l}|\Phi(\mathbf{0}, \mathbf{0})\rangle = |\Phi_{kl}^{T2}\rangle$. Note that this term is only nonzero if x^k and x^l correspond to different sites of the MPS chain. This part corresponds to two single excitations on different sites. Also note that $\frac{\partial^2}{\partial \overline{x_k} \partial \overline{x_l}}|\Phi(\mathbf{0}, \mathbf{0})\rangle = 0$.

4. The local part of the double tangent space $\frac{\partial^2}{\partial \overline{x_k} \partial x^l}|\Phi(\mathbf{0}, \mathbf{0})\rangle = |\Phi_{\overline{k}l}^{T2}\rangle$. Note that this term is only nonzero if $\overline{x_k}$ and x^l belong to the same site, and correspond to the same row index in the matrix notation $x[i]$. This part consists of all possible connections between the renormalized basis states $|L_\alpha^i\rangle$ (from $|L_\alpha^{i-1}\rangle \otimes |n_i\rangle$) and $|R_\beta^i\rangle$.

These states are not all mutually orthogonal. Note that the local part of the double tangent space arises because we have considered a unitary variant of the Thouless theorem for MPS. The original (non-unitary) Thouless parametrization for HF depends only on the complex parameters, and not on their complex conjugates.

If two excitation operators in HF try to annihilate an occupied single particle twice, the state is destroyed. The space of double OV excitations therefore consists of the replacement of two different occupied single particles by two different virtual single particles.

The local part of the double tangent space of an MPS can be written as $B^\mu |\Phi_\mu^0\rangle$, which lies entirely in the space spanned by the MPS reference $|\Phi^0\rangle$ and the non-redundant tangent space vectors $|\Phi_k^T\rangle$. Together with the other two arguments above, this provides a justification to discard this part of the double tangent space without any loss in variational freedom, and to consider only two single excitations acting on different sites, for the double tangent space.

D. Excitation operators

The excitation operators for an MPS can be read from the Thouless theorem:

$$|\Phi_k^T\rangle = \hat{B}_k^\dagger |\Phi^0\rangle = \left. \frac{\partial}{\partial x^k} |\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle \right|_{\mathbf{x}=\mathbf{0}}. \qquad (52)$$

See e.g. Sec. IV in Rowe *et al.*[13] for a discussion on the relationship between the linearized time-dependent variational principle on a general manifold, and the EOM approach to the RPA equations. The operators $\hat{B}_k^\dagger$ are obtained by going from the manifold representation based on the virtual space in Eq. (46), to a representation based on the physical Hilbert space $|\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle = \exp\left(x^k \hat{B}_k^\dagger - \overline{x_k}\hat{B}^k\right)|\Phi^0\rangle$. When only the first order tangent space needs to match, $\hat{B}_k^\dagger = |\Phi_k^T\rangle\langle\overline{\Phi^0}|$ can be used. It will be a challenge to find the $\hat{B}_k^\dagger$'s to match the higher order tangent spaces too. Finding an answer to this problem, is closely related to finding a site-space analogue of the particle Fock space, based on the MPS reference.

From Eq. (46), it can be understood that this excitation operator projects out the site-matrices $A_0[i(k)]$ and replaces them with the tangent space site-matrices $\partial_{x_k} B(\mathbf{x})[i(k)]|_{\mathbf{x}=\mathbf{0}}$. It adds a single excitation to the vacuum $|\Phi^0\rangle$. In the chosen gauge, a single MPS excitation is localized to one site, just like a single OV excitation of an SD is localized to one orbital.

From Eq. (46), it can also be understood that a deexcitation operator projects out the tangent space site-matrices $\partial_{x_k} B(\mathbf{x})[i(k)]|_{\mathbf{x}=\mathbf{0}}$ and replaces them with the site-matrices $A_0[i(k)]$. Remember that the tangent space metric is the unit matrix for the chosen parametrization, and that the deexcitation projections are hence not only orthogonal to the MPS reference (they destroy the vacuum $|\Phi^0\rangle$), but also orthogonal to other tangent space site-matrices:

$$\hat{B}_l |\Phi^0\rangle = 0 \qquad (53)$$
$$\hat{B}_l |\Phi_k^T\rangle = \delta_{l,k} |\Phi^0\rangle. \qquad (54)$$

The deexcitation operators of the ket vectors, are the excitation operators of the bra vectors:

$$\langle\overline{\Phi^0}| \hat{B}_l = \langle\Phi_l^T| \qquad (55)$$

Consider the commutators $\left[\hat{B}_l^\dagger, \hat{B}_k^\dagger\right]$ and $\left[\hat{B}_l, \hat{B}_k^\dagger\right]$. Their general expressions are far from trivial, only their expectation value with respect to the vacuum $|\Phi^0\rangle$ is clear:

$$\langle\overline{\Phi^0}| \left[\hat{B}_l^\dagger, \hat{B}_k^\dagger\right] |\Phi^0\rangle = 0 \qquad (56)$$
$$\langle\overline{\Phi^0}| \left[\hat{B}_l, \hat{B}_k^\dagger\right] |\Phi^0\rangle = \delta_{k,l}. \qquad (57)$$

A bosonic algebra for the excitation operators is hence only retrieved when expectation values with respect to the vacuum are taken. The operators $\hat{B}_k^\dagger$ are called quasi-boson operators.

VII. OPTIMAL TIME EVOLUTION FOR MPS

The optimal time evolution of an MPS, which stays within the MPS ansatz space, was derived by means of the TDVP in Ref. 23 and 21. Now that we have established the Thouless theorem for MPS, we can rephrase the result as

$$i\hbar \dot{x}^k(\mathbf{x}=\mathbf{0}) = \langle\overline{\Phi^0}| \hat{B}^k \hat{H} |\Phi^0\rangle. \qquad (58)$$

Also in this case, Eq. (58) can be obtained by inserting $|\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle$ in the time-dependent Schrödinger equation, and by projecting the time-dependent equation onto $\hat{B}_k^\dagger |\Phi(\mathbf{x},\overline{\mathbf{x}})\rangle$:

$$i\hbar \langle\overline{\Phi}| \hat{B}^k \left(\dot{x}^l \frac{\partial}{\partial x^l} + \dot{\overline{x}}_l \frac{\partial}{\partial \overline{x_l}}\right) |\Phi\rangle = \langle\overline{\Phi}| \hat{B}^k \left(\hat{H} - E_{\mathrm{MPS}}\right) |\Phi\rangle \qquad (59)$$

Evaluation for $\mathbf{x}=\mathbf{0}$ yields Eq. (58). This form of time propagation automatically stays within the MPS ansatz space. No Hamiltonian decompositions or bond dimension truncations are necessary.[21,23]

VIII. RPA FOR MPS

A. In a redundant parametrization

One way to obtain the RPA equations for MPS, is to consider the linearized time-dependent equations in the vicinity of a variational minimum, and to project them onto the tangent space of the manifold.[21,25] Consider a small time-dependent step around the minimum $A^\mu(t) = A_0^\mu + B^\mu(t)$. The time-dependent equation, its projection onto the tangent space, and its first order terms become:

$$i\hbar |\Phi_\nu(\mathbf{A})\rangle \dot{A}^\nu = \left(\hat{H} - E_{\mathrm{MPS}}\right) |\Phi(\mathbf{A})\rangle \qquad (60)$$

$$i\hbar \langle\overline{\Phi_\mu(\mathbf{A})}| \Phi_\nu(\mathbf{A})\rangle \dot{A}^\nu = \langle\overline{\Phi_\mu(\mathbf{A})}| \hat{H} - E_{\mathrm{MPS}} |\Phi(\mathbf{A})\rangle \qquad (61)$$

$$i\hbar \langle\overline{\Phi_\mu(\mathbf{A}_0)}| \Phi_\nu(\mathbf{A}_0)\rangle \dot{B}^\nu = \langle\overline{\Phi_\mu(\mathbf{A}_0)}| \hat{H} - E_{\mathrm{MPS}} |\Phi_\nu(\mathbf{A}_0)\rangle B^\nu + \langle\overline{\Phi_{\mu\nu}(\mathbf{A}_0)}| \hat{H} - E_{\mathrm{MPS}} |\Phi(\mathbf{A}_0)\rangle \overline{B}^\nu. \qquad (62)$$

with $E_{\mathrm{MPS}} = \langle\overline{\Phi^0}| \hat{H} |\Phi^0\rangle$. By taking a harmonic ansatz for the perturbation $\mathbf{B} = \mathbf{Y}e^{-i\omega t} + \overline{\mathbf{Z}}e^{i\omega t}$, the RPA equa-

tions are found:

$$\hbar\omega \begin{bmatrix} S & 0 \\ 0 & -\overline{S} \end{bmatrix} \begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} = \begin{bmatrix} H & W \\ \overline{W} & \overline{H} \end{bmatrix} \begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} \qquad (63)$$

with $S_{\mu\nu} = \langle \overline{\Phi_\mu^0} \mid \Phi_\nu^0 \rangle$, $H_{\mu\nu} = \langle \overline{\Phi_\mu^0} \mid \hat{H} - E_{\mathrm{MPS}} \mid \Phi_\nu^0 \rangle$, and $W_{\mu\nu} = \langle \overline{\Phi_{\mu\nu}^0} \mid \hat{H} - E_{\mathrm{MPS}} \mid \Phi^0 \rangle$. Note that $\| \left(\hat{H} - E_{\mathrm{MPS}} \right) \mid \Phi^0 \rangle \|_2$ becomes smaller when $\mid \Phi^0 \rangle$ becomes a better approximation for the exact ground state. $\|W\|_2$ is hence a measure for the accuracy of the MPS approximation to the exact ground state.[25] A specific eigenvector of Eq. (63) can be obtained in $\mathcal{O}(LD^3)$ complexity.[21,28]

B. In a non-redundant parametrization

By changing the basis from $\mid \Phi_\mu^0 \rangle$ to $\mid \Phi_k^T \rangle = Z_k^\mu \mid \Phi_\mu^0 \rangle$, with $\frac{\partial}{\partial x^k} B^\mu(\mathbf{x}) = Z_k^\mu$, the overlap matrix S becomes the unit matrix I: $\overline{Z}_k^\mu S_{\mu\nu} Z_l^\nu = \delta_{kl}$. Analogously, the hermitian matrix A and the complex symmetric matrix B, both of dimension $\dim\mathbb{T} \times \dim\mathbb{T}$, can be defined as resp. $A_{kl} = \overline{Z}_k^\mu H_{\mu\nu} Z_l^\nu$ and $B_{kl} = \overline{Z}_k^\mu W_{\mu\nu} Z_l^\nu$. The RPA equations become:

$$\hbar\omega \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} \qquad (64)$$

where $\mathbf{y}$ and $\mathbf{z}$ are now coefficients with respect to $\mid \Phi_k^T \rangle$. The same result can be obtained by linearizing Eq. (59), just as for HF.

The A- and B-matrices can be constructed explicitly. If only a few excitation energies are desired, it is better to resort to a sweep algorithm, which can be implemented in an existing DMRG code.[28] Implementation details of this sweep algorithm will be presented elsewhere.

C. EOM derivation

The excitation operators discussed in Sec. VI D allow for a rederivation of the RPA equations for MPS by means of the EOM. An exact bosonic algebra can be set up by adding correction terms to operators defined in Sec. VI D, so that $\left[\hat{B}_l^\dagger, \hat{B}_k^\dagger \right] = 0$ and $\left[\hat{B}^l, \hat{B}_k^\dagger \right] = \delta_{l,k}$. A justification is given by Eqs. (56) and (57). The Hamiltonian can be expanded in these bosonic operators, and RPA coincides with truncating the expansion after second order. Expressions for the RPA correlation energy and wave-function follow, just as for HF:

$$E_{\mathrm{cRPA}} = -\frac{\mathrm{tr}A}{2} + \sum_{\omega_n > 0} \frac{\hbar\omega_n}{2} = -\sum_n \hbar\omega_n \sum_k \mid z_n^k \mid^2 \quad (65)$$

$$\mid \mathrm{RPA} \rangle \propto e^{\frac{1}{2}\left(\overline{zy^{-1}} \right)^{k;l} \hat{B}_k^\dagger \hat{B}_l^\dagger} \mid \Phi^0 \rangle . \qquad (66)$$

IX. POST-DMRG METHODS

A. TDA and Brillouin's theorem

A preferred tangent basis can be found by searching the eigenstates of the Hamiltonian in the basis $\mid \Phi_k^T \rangle$. This corresponds to diagonalizing the A-matrix of the RPA Eq. (64). As $\mid \Phi^0 \rangle = \frac{1}{L} A_0^\mu \mid \Phi_\mu^0 \rangle$ gave the lowest energy solution for the ansatz $B^\mu \mid \Phi_\mu^0 \rangle$ and $\mid \Phi_k^T \rangle \perp \mid \Phi^0 \rangle$, approximations for excited states are found this way. This is the MPS analogue of TDA.[4,5,29] For a variational minimum

$$0 = \frac{\partial E}{\partial \overline{A_0^\mu}} = \frac{\langle \overline{\Phi_\mu^0} \mid \hat{H} \mid \Phi^0 \rangle}{\langle \overline{\Phi^0} \mid \Phi^0 \rangle} - \frac{\langle \overline{\Phi^0} \mid \hat{H} \mid \Phi^0 \rangle}{\langle \overline{\Phi^0} \mid \Phi^0 \rangle^2} \langle \overline{\Phi_\mu^0} \mid \Phi^0 \rangle .$$

$$(67)$$

If the wave-function $\mid \Phi^0 \rangle$ is normalized and only norm- and phase-conserving changes $\hat{B}_k^\dagger \mid \Phi^0 \rangle \perp \mid \Phi^0 \rangle$ are considered,[23,28]

$$\langle \overline{\Phi_k^T} \mid \hat{H} \mid \Phi^0 \rangle = \langle \overline{\Phi^0} \mid \hat{B}^k \hat{H} \mid \Phi^0 \rangle = 0. \qquad (68)$$

This is the MPS analogue of Brillouin's theorem.[4,30] For MPS, excited momentum eigenstates of translationally invariant systems have previously been approximated in the non-redundant tangent space basis.[38]

B. CC and CI

The Thouless theorem for MPS and Eq. (66) suggest CC and CI ansatzes on top of an MPS reference. Consider for example the single and double excitations:

$$\mid \mathrm{CCSD} \rangle \propto e^{y^k \hat{B}_k^\dagger + \frac{1}{2} z^{kl} \hat{B}_k^\dagger \hat{B}_l^\dagger} \mid \Phi^0 \rangle \qquad (69)$$

$$\mid \mathrm{CISD} \rangle \propto \left(x + y^k \hat{B}_k^\dagger + \frac{1}{2} z^{kl} \hat{B}_k^\dagger \hat{B}_l^\dagger \right) \mid \Phi^0 \rangle . \qquad (70)$$

With the exposition in Sec. V, VI C and VI D, we can also propose the following CCSD and CISD ansatzes:

$$\mid \mathrm{CCSD} \rangle \propto e^{C^{\mu\nu} \partial_\mu \partial_\nu} \mid \Phi^0 \rangle \qquad (71)$$

$$\mid \mathrm{CISD} \rangle \propto C^{\mu\nu} \mid \Phi_{\mu\nu}^0 \rangle \qquad (72)$$

with $\mathcal{O}(\frac{1}{2}(LdD^2)^2)$ parameters in the symmetric C-tensor. Note that working in the Lth order tangent space corresponds to the FCI ansatz. For DMRG (HF) the CCSD and CISD ansatzes can be considered as a way to improve the correlation between two sites (electrons) embedded in an approximate environment given by the zeroth order MPS (SD). Since the double tangent space can connect sites that are far apart, this enables the CCSD and CISD expressions to directly build in long-range entanglement. If the zeroth order description fails (static correlation for HF, critical system for DMRG), these ansatzes will fail too. Also for MPS, CISD is not size-consistent if there are more than two sites in the compound system, whereas CCSD is always size-consistent because of the exponential ansatz.

X. SYMMETRY-ADAPTED CALCULATIONS

For large calculations, symmetry-adapted MPS ansatzes are often used. They allow to search for eigenstates within a specific symmetry sector of the total

Hilbert space, and lead to computational advantages in memory and time. An MPS ansatz without symmetry-adaptation can yield an approximate eigenstate which breaks the symmetry. Its tangent space then also contains symmetry-broken vectors. RPA-MPS breaks down if a symmetry multiplet of a non-Abelian group is incomplete at a certain MPS boundary. Therefore we use symmetry-adapted MPS ansatzes for the applications.

A. Tangent space without symmetry adaptation

First consider an MPS ansatz without symmetry adaptation. A basis for its non-redundant tangent space, which is at the same time a basis of symmetry eigenvectors, can only be constructed when the MPS reference is an eigenvector of those symmetries. If the MPS reference is a symmetry eigenvector, its tangent space (in general) also contains symmetry eigenvectors which belong to a different irreducible representation (irrep). We provide a simple counting argument.

Consider an MPS with length L even, then the center virtual dimension has to be $d^{\frac{L}{2}}$ to represent a general FCI state.[34] The number of states in its non-redundant tangent space is then

$$\dim \mathbb{T} = \sum_{k=1}^{\frac{L}{2}} \left(dd^k - d^{k-1} \right) d^{k-1} = d^L - 1, \qquad (73)$$

i.e. the rest of the Hilbert space. Note that these $d^L - 1$ non-redundant tangent space vectors can only be constructed if all Schmidt values are greater than zero.[23,32] Suppose that this is the case. The MPS reference and its non-redundant tangent space then span the entire Hilbert space. If the MPS reference transforms according to a particular irrep of the symmetry group of the Hamiltonian, a basis of symmetry eigenvectors can be constructed for its non-redundant tangent space. If the Hilbert space is spanned by symmetry vectors belonging to at least two different irreps, the non-redundant tangent basis then contains symmetry eigenvectors belonging to a different irrep than the MPS reference.

B. Implications for RPA

If an MPS ansatz without symmetry adaptation is variationally optimized, it can occur that due to the choice of virtual dimensions a symmetry multiplet of a non-Abelian group (e.g. SU(2)) is incomplete at a certain boundary. From the projector interpretation of the non-redundant tangent space, it can be understood that this may lead to spurious zero energy RPA excitations: replacing an occuring renormalized basis state of the multiplet by one that is not in the renormalized basis, can lead to a state with the same energy and hence a spurious zero energy RPA excitation. For this reason, we have

opted to use symmetry-adapted MPS references in this work.

C. Tangent space of a symmetry-adapted ansatz

We now discuss the construction of the tangent space of an SU(2) ⊗ U(1) adapted MPS ansatz. A spin- and particle number-adapted MPS decomposes into Clebsch-Gordan coefficients of the imposed symmetry groups and reduced tensors, due to the Wigner-Eckart theorem:[18,39]

$$A^{ss^z N}_{j_L j_L^z N_L \alpha_L; j_R j_R^z N_R \alpha_R}$$
$$= \langle j_L j_L^z ss^z \mid j_R j_R^z \rangle \, \delta_{N_L+N,N_R} T^{sN}_{j_L N_L \alpha_L; j_R N_R \alpha_R} \quad (74)$$

The derivative operator $\frac{\partial}{\partial A^\mu}$ in Eq. (33) is then replaced by $\frac{\partial}{\partial T^\kappa}$. All symmetry imposing Clebsch-Gordan coefficients are still in place, and the tangent space vectors hence have the same symmetry as the MPS reference. The non-redundant tangent space can be constructed in an analogous way as for the case without symmetry adaptation. The entire symmetry sector of the Hilbert space (minus the MPS reference) is retrieved in the non-redundant tangent space, if the virtual dimensions are taken as large as possible. The difference between the tangent spaces with and without symmetry-adaptation can be compared to the restricted and unrestricted HF manifolds.[13] For the former only singlet excitations are possible, while for the latter triplet excitations are allowed too, even if the ground state is a singlet.

Note that if a symmetry-adapted MPS is optimized by the imaginary time-evolution of Sec. VII, the distribution of the bond dimensions over the symmetry sectors is fixed. As such an optimization does not lead to an optimal distribution of the bond dimensions, we have used the two-site DMRG algorithm to optimize all the MPS reference wave-functions in this work.

Henceforth symmetry-adapted will be used as a shorthand for spin- and particle number-adapted.

XI. THE 1D HUBBARD CHAIN

In this section, we approximate low-lying eigenstates of the one-dimensional Hubbard chain with open boundary conditions (OBC) (see Eq. (34)). The CISD-MPS ansatz of Eq. (72), which contains all excitations to the double tangent space, is used to improve on the ground state and to find low-lying excitations. The results are compared with TDA-MPS, which contains all excitations to the single tangent space. With RPA-MPS, we search for the Goldstone mode of a symmetry-broken ground state. In addition, we discuss RPA-MPS correlation energy calculations.

TABLE I. CISD-MPS improvement on the ground state and low-lying excitation energies for the one-dimensional Hubbard chain with length $L = 8$ and OBC. D is the number of SU(2) multiplets retained at each boundary in the symmetry-adapted MPS reference calculation. The state for which the absolute energy is shown, was chosen as MPS reference.
$^{(a)}$ Excitation with different multiplicity. The required FCI excitation is not in the TDA-MPS spectrum.

U	Quantity	S	N	Exact (FCI)	TDA-MPS [D=3]	TDA-MPS [D=9]	CISD-MPS [D=3]
0.1	E_0	0	8	-9.319312	-9.067465	-9.301264	-9.315185
	E_1-E_0	$\frac{1}{2}$	7	0.297631	0.222150	0.285466	0.311181
	E_2-E_0	$\frac{1}{2}$	9	0.397631	0.322150	0.385466	0.411181
	E_3-E_0	0	6	0.611620	0.873285	0.619417	0.629720
1	E_0-E_1	$\frac{1}{2}$	7	-0.022354	-0.082237	-0.029340	0.011799
	E_1	0	6	-7.790647	-7.532068	-7.780764	-7.785715
	E_2-E_1	0	8	0.095814	$0.543100^{(a)}$	0.105942	0.135785
	E_3-E_1	1	6	0.517393	$0.572255^{(a)}$	0.530944	0.542513
10	E_0	0	4	-5.187427	-5.083270	-5.186955	-5.187090
	E_1-E_0	$\frac{1}{2}$	5	0.008950	-0.010314	0.009270	0.010988
	E_2-E_0	1	4	0.113988	0.127636	0.114721	0.114984
	E_3-E_0	$\frac{1}{2}$	5	0.189005	$0.205577^{(a)}$	0.196828	0.192880
100	E_0	0	4	-4.805753	-4.736845	-4.805615	-4.805360
	E_1-E_0	1	4	0.013020	0.013672	0.013016	0.013783
	E_2-E_0	1	4	0.027045	0.022700	0.027013	0.028034
	E_3-E_0	0	4	0.034327	$0.316707^{(a)}$	0.034327	0.035940

A.   CISD-MPS

The TDA and CISD calculations were done by optimizing a symmetry-adapted MPS reference, with D retained multiplets at each boundary. This reference was then used in TDA and CISD calculations without symmetry constraints. As the symmetry-adapted reference is not necessarily a variational minimum for an MPS ansatz without symmetry constraints, negative excitation energies can occur.

The CISD ansatz in Eq. (72) leads to a generalized eigenvalue problem

$$\langle \overline{\Phi^0_{\kappa\lambda}} \mid \hat{H} \mid \Phi^0_{\mu\nu} \rangle C^{\mu\nu} = E \langle \overline{\Phi^0_{\kappa\lambda}} \mid \Phi^0_{\mu\nu} \rangle C^{\mu\nu} \quad (75)$$

which was solved by multi-targeting the lowest states with the Davidson algorithm.[40] By decomposing the C-tensor, the CISD ansatz can be written as a sum over MPS wave-functions:

$$|\text{CISD}\rangle = \sum_{i<j} C^{\mu_i\nu_j} |\Phi^0_{\mu_i\nu_j}\rangle = \sum_{i<j}\sum_{p(i,j)} C^{\mu_i}_{L,p} C^{\nu_j}_{R,p} |\Phi^0_{\mu_i\nu_j}\rangle. \quad (76)$$

This allows to use standard MPS machinery[34] in the matrix-vector multiplication. Because the sum of several MPS wave-functions yields an MPS with a larger bond dimension,[34] this immediately leads to the understanding that the CISD ansatz can introduce extra entanglement. We chose $L = 8$ and four U-values: 0.1, 1, 10 and 100. With increasing U, the ground state changes from a collection of quasi-free particles to a highly correlated state. For the latter, HF gives a qualitatively wrong description. For $U = 1$, the ground state contains 7 particles

and has spin $\frac{1}{2}$. If a symmetry-broken reference is chosen, the multiplet degeneracy of the excitations is lost. Therefore, we opted for the first singlet state as MPS reference for $U = 1$. Although the TDA and CISD calculations were not symmetry-adapted, the multiplet degeneracy was *exactly* retrieved, because we started from a symmetry-adapted MPS reference. The first four multiplets for each U-value are shown in Table I.

The TDA-MPS [D=9] energies and the CISD-MPS [D=3] energies are of roughly the same quality, and improve on the TDA-MPS [D=3] energies significantly. The CISD-MPS ansatz contains $\mathcal{O}(\frac{1}{2}(LdD^2)^2)$ variational parameters to include extra correlation between all pairs of sites on top of the MPS reference, and can be used both to improve on the ground state and to approximate excited states. The TDA-MPS ansatz contains $\mathcal{O}(LdD^2)$ variational parameters, and due to the MPS analogue of Brillouin's theorem, it can only be used to approximate excited states. The relative accuracy of the CISD-MPS [D=3] and MPS [D=9] reference state energies changes with U. With increasing single-particle nature, the CISD-MPS ansatz performs better for the targeted reference.

For small D, not all excited states are retrieved with TDA-MPS. An example is the third excited state for D=3 and U=100. The targeted state consists of two singlet-triplet excitations, which interact to form a bound singlet state. This is well captured by CISD-MPS [D=3] and TDA-MPS [D=9], while E_3 for TDA-MPS [D=3] in Table I corresponds in fact to a doublet. The MPS [D=9] reference has a large enough bond dimension to capture the two excitations in its single tangent space, while the CISD-MPS [D=3] ansatz captures the double excitation
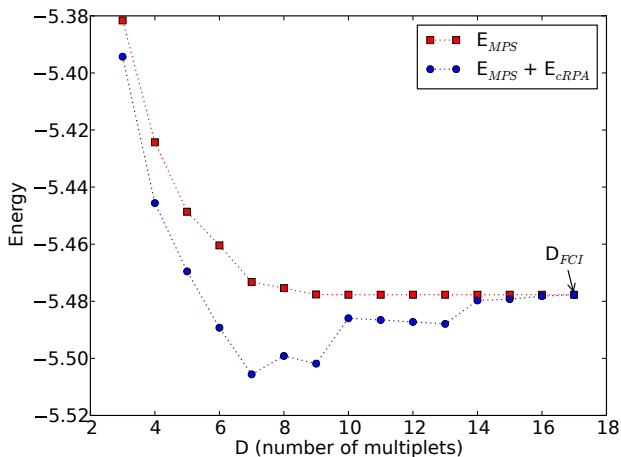
FIG. 1. The RPA-MPS correlation energy for the Hubbard chain with OBC, length $L = 6$, filled with $N = 4$ electrons, in the singlet state, and $U = 1$. A symmetry-adapted ansatz was used, with D retained multiplets at each boundary.

in the MPS's double tangent space.

B. RPA-MPS and Goldstone modes

The $L = 8$ and $U = 1$ case is an ideal candidate to retrieve a Goldstone mode, because a specific spin-$\frac{1}{2}$ ground state vector is necessarily a symmetry-broken state. With an MPS reference optimized without any symmetry constraints and $D = 16$ (now exceptionally the number of states instead of the number of multiplets), we find one zero energy solution to the RPA equations, and this solution also satisfies Eq. (25). This is the Goldstone mode from the symmetry-broken spin-$\frac{1}{2}$ ground state. Zero energy solutions can also arise for singlet ground states, if the MPS accidently breaks non-Abelian symmetries, as discussed in Sec. X. This can be avoided by retaining complete multiplets at each boundary, whereas the RPA Goldstone mode for a symmetry-broken ground state will always occur, even for bond dimensions that reproduce the full Hilbert space.

### C. The RPA-MPS correlation energy

We calculated RPA-MPS correlation energies for symmetry-adapted ansatzes. Remember that only excitations with the same symmetry as the MPS reference are then retrieved. For the Hubbard chain with OBC and length $L = 6$, filled with $N = 4$ electrons, in the singlet state, and $U = 1$, E_{cRPA} is shown in Fig. 1. From Eq. (65), $2E_{\mathrm{cRPA}}$ can be interpreted as the mean difference between RPA and TDA excitation energies, multiplied by the number of excitations (dim$\mathbb{T}$). With increasing D, $|E_{\mathrm{cRPA}}|$ first increases because the number of excitations increases. For even larger D, the mean difference between the RPA and TDA energies vanishes

faster than the number of excitations increases. When the FCI virtual dimensions are reached, the RPA and TDA excitation energies are equal, as the B-matrix vanishes, and E_{cRPA} is exactly zero.

When calculating E_{cRPA}, care has to be taken that the MPS reference is the variational minimum and that no symmetries are broken, such as the multiplet structure at a boundary or e.g. the SO(4) symmetry when considering a half-filled Hubbard chain.[41] If these conditions are not fulfilled, the RPA-MPS correlation energy breaks down ($|E_{\mathrm{cRPA}}| \gg |E_{\mathrm{MPS}}|$).

XII. POLYENES

Polyenes are linear conjugated chains of hydrocarbons:

$$CH_2 = CH - CH = CH - CH = CH_2. \tag{77}$$
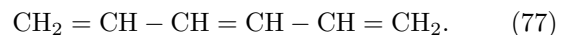
Excitations in the π-system lie in the visible region of the spectrum, and polyenes are therefore important building blocks for light absorption and dyes. They have a long history of use as benchmark systems to test new quantum chemistry excited state methods. The π-system can be approximated by the long-range Pariser-Parr-Pople (PPP) Hamiltonian, where the two-body terms of the Hamiltonian in Eq. (5) are approximated by a local Coulomb repulsion:

$$\hat{V} = \frac{1}{2} \sum_{kl\sigma\tau} R^{kl} \hat{n}_{k\sigma} \hat{n}_{l\tau}. \tag{78}$$

The Latin letters denote orbitals and the Greek letters spin projections. For our calculations, we used the Ohno parametrization for the electron-electron repulsion R^{kl}.[42] All Hamiltonian parameters, except $R_{\mathrm{single}\atop\mathrm{double}} = 1.40 \pm 0.05 \mathring{A}$, are identical to the ones in Ref. 43.

Many DMRG calculations studying the excited states and response properties of conjugated molecules have been performed, using a parametrized Hamiltonian.[44] At the ab initio level, high-lying excited states have been targeted with the harmonic Davidson adaptation of the DMRG method.[45] Frequency-dependent dipole polarizabilities were computed at the ab initio level by Dorando et al.[28] using the TDA-MPS approximation.

Using the PPP Hamiltonian, we approximated the first three particle-conserving singlet excitations with the symmetry-adapted RPA-MPS and TDA-MPS methods. We kept $D = 20$ retained multiplets at each boundary. The TDA-MPS excitation energies are shown in Fig. 2 as a function of the number of carbon atoms N in the polyene. The symmetry labeling was based on Fig. 10 in Ref. 46. The difference between the RPA-MPS and TDA-MPS energies is shown in Fig. 3, indicating that the ground state MPS reference is already quite accurate for $D = 20$, as the B-matrix contributions are small. The RPA-MPS and TDA-MPS excitation energies match better for the higher excitations of Fig. 2.
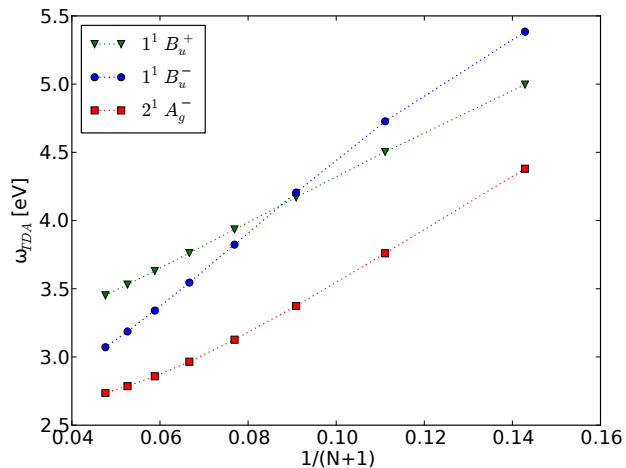
FIG. 2. The first three TDA-MPS excitation energies for a polyene chain with N carbon atoms, for which the π-system was approximated by the long-range PPP Hamiltonian.

XIII. SUMMARY

In this work, we attempted to set up a post-DMRG framework by finding the excitation structure of the MPS reference. As a guide, we carefully followed the structure of HF theory and the subsequent post-HF methods, exploiting the fact that both HF and DMRG can be seen as product-like wave-functions.[20]

A variational wave-function ansatz can be used in the TIVP to yield self-consistent equations. With the TDVP, optimal time-evolution is found which stays within the ansatz manifold. Linearization of the TDVP around a variational minimum gives the RPA equations. The optimal time-evolution requires a non-redundant parametrization of the ansatz's tangent space to exclude
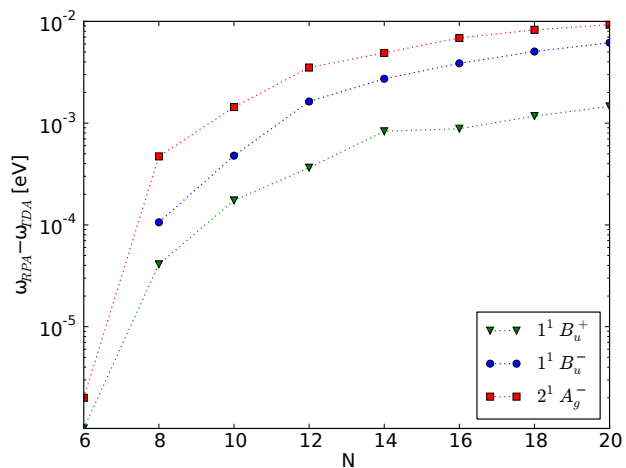

FIG. 3. The difference between the RPA-MPS and TDA-MPS energies for the first three excitations of a polyene chain with N carbon atoms, for which the π-system was approximated by the long-range PPP Hamiltonian.

meaningless variations of the wave-function. Occupied-occupied variations in HF theory, as well as variations in the direction of the renormalized DMRG basis states, only lead to norm or phase changes. They do not change the physical state represented by the ansatz.

Exponentiation of the norm- and phase-conserving variations in HF theory, led to the Thouless theorem: a non-redundant parametrization of the entire HF (Grassmann) manifold, generated by the OV excitations of any particular SD. In this work, we have proposed the DMRG counterpart: a non-redundant parametrization of the entire MPS manifold, generated by the norm- and phase-conserving changes of any particular MPS wave-function. Just like the norm- and phase-conserving changes of HF theory are generated by replacing an occupied orbital by a virtual orbital, the norm- and phase-conserving changes of an MPS wave-function are generated by replacing the occuring renormalized basis states by discarded renormalized basis states. We have proven the MPS counterpart of Thouless's theorem for a general MPS with OBC, for which no Schmidt values vanish.

By identifying the excitation structure of the SD/MPS ansatz by means of the Thouless theorem, the RPA equations can be rederived be means of the EOM. This allows for a bosonic expansion of the Hamiltonian, and the definition of the RPA correlation energy and wave-function.

The different orders of tangent space of the Thouless parametrization generate the CI basis. Eigenstates of the Hamiltonian can be approximated in this basis. CIS, or CI with only single excitations, yields again the SD/MPS reference due to Brillouin's theorem, as well as a set of excited states. These excited states are found by diagonalizing the Hamiltonian in the non-redundant tangent basis, or the A-matrix of RPA. This method is known as TDA.

When the MPS reference is a good approximation of the true ground state, $\| \left(\hat{H} - E_{\mathrm{MPS}} \right) |\Phi^0\rangle \|_2$ becomes small, and the B-matrix contributions of RPA vanish. TDA and RPA then lead to the same excitation energies.

The RPA wave-function suggests a size-consistent CC ansatz on top of the reference wave-function.

The ideas presented in this paper are illustrated with proof-of-principle calculations of CISD-MPS improvements on the ground state, TDA-MPS, RPA-MPS and CISD-MPS excitation energies, an RPA-MPS Goldstone mode, and the RPA-MPS correlation energy. In contrast to HF, the MPS reference gives also in the highly correlated regime of the Hubbard model a qualitatively good description, and variational post-DMRG methods such as TDA-MPS and CISD-MPS give numerically relevant results. For an MPS with small bond dimensions, two correlated single excitations are not always retrieved in the tangent space, and the CISD-MPS ansatz is a better choice then.

Near completion of this work, we learned about Ref.

47 which presents RPA-MPS calculations and new multi-site excitation ansatzes for uniform MPS.

Appendix: Explicit Grassmann manifold parametrization

The proof given here is inspired by the proof for the unitary counterpart of Thouless's theorem for HF, given in Rowe *et al.*[13] Give a unitary $m \times n$ matrix Q with $m > n$, i.e. $Q^\dagger Q = I_n$, and a second unitary matrix U, of the same form as Q. Form a unitary $m \times (m-n)$ matrix $\tilde{Q}$ so that $[Q\tilde{Q}][Q\tilde{Q}]^\dagger = I = [Q\tilde{Q}]^\dagger[Q\tilde{Q}]$. For the parametrization

$$\tilde{A}(\mathbf{y}, \overline{\mathbf{y}}) = \exp\left(\tilde{Q}yQ^\dagger - Qy^\dagger\tilde{Q}^\dagger\right)Q \qquad (A.1)$$

with y an $(m-n) \times n$ matrix containing the complex variables and $\mathbf{y}$ the corresponding flattened column, there exists at least one $\mathbf{y_u}$ so that the columns of $\tilde{A}(\mathbf{y_u}, \overline{\mathbf{y_u}})$ and the columns of U (denoted by $\mathbf{u_k}$) span the same space. We will provide a proof by construction:

1. The matrix $M = U^\dagger QQ^\dagger U$ is a hermitian positive semidefinite matrix. There exists a unitary transformation to rotate the basis $\mathbf{u_i}$ to $\mathbf{v_i}$, so that $\mathbf{v_k}^\dagger QQ^\dagger \mathbf{v_j} = \delta_{kj}n_k^2$.

2. Write $\mathbf{v_i}$ in terms of $\mathbf{q_k}$ and $\tilde{\mathbf{q}}_\mathbf{k}$, the columns of Q and $\tilde{Q}$: $\mathbf{v_i} = \sum_k \alpha_{ik}\mathbf{q_k} + \sum_l \beta_{il}\tilde{\mathbf{q}}_\mathbf{l}$. From the previous step we know that $\delta_{ij} = \mathbf{v_i}^\dagger \mathbf{v_j} = n_i^2\delta_{ij} + \sum_l \overline{\beta_{il}}\beta_{jl}$.

3. If $n_i \neq 0$, define $\mathbf{r_i}$ by $n_i\mathbf{r_i} = \sum_k \alpha_{ik}\mathbf{q_k}$. If $n_i \neq 1$, define $\tilde{\mathbf{r}}_\mathbf{i}$ by $(1-n_i^2)^{\frac{1}{2}}\tilde{\mathbf{r}}_\mathbf{i} = \sum_l \beta_{il}\tilde{\mathbf{q}}_\mathbf{l}$. Note that if e.g. $2n > m$, a number of n_i will certainly be 1, and the corresponding vectors $\tilde{\mathbf{r}}_\mathbf{i}$ cannot be constructed.

4. From the previous steps, it follows that the vectors $\{\mathbf{r_i}, \tilde{\mathbf{r}}_\mathbf{i}\}$ are orthonormal. Complete both sets with additional vectors, so that they span the same space as the columns of resp. Q and $\tilde{Q}$. If the matrix R contains $\mathbf{r_i}$ in its columns and $\tilde{R}$ contains $\tilde{\mathbf{r}}_\mathbf{i}$ in its columns, a unitary transformation P links R and Q by $R = QP$ and a unitary transformation $\tilde{P}$ links $\tilde{R}$ and $\tilde{Q}$ by $\tilde{R} = \tilde{Q}\tilde{P}$.

5. If $n_i \neq 1$, consider $\mathbf{w_i} = \exp\left(\gamma_i(\tilde{\mathbf{r}}_\mathbf{i}\mathbf{r_i}^\dagger - \mathbf{r_i}\tilde{\mathbf{r}}_\mathbf{i}^\dagger)\right)\mathbf{r_i} = \cos(\gamma_i)\mathbf{r_i} + \sin(\gamma_i)\tilde{\mathbf{r}}_\mathbf{i}$. Assign $0 \leq \gamma_i \leq \frac{\pi}{2}$ so that $\cos(\gamma_i) = n_i$. It then follows that $\mathbf{w_i} = \mathbf{v_i}$. Note that if $n_i = 1$, γ_i would have been 0, and that the exponential in front of $\mathbf{r_i}$ then becomes the identity. So it poses no problem that the corresponding vectors $\tilde{\mathbf{r}}_\mathbf{i}$ cannot be constructed.

6. If γ is regarded as a diagonal matrix containing the γ_i values, the singular value decomposition of y_u is given by $y_u = \tilde{P}\gamma P^\dagger$. This can be confirmed by writing the exponential expression for $\mathbf{w_i}$ in terms of Q and $\tilde{Q}$.

This concludes the construction of the complex $(m-n) \times n$ matrix y_u. Eq. (A.1) hence represents a Grassmann manifold.

* sebastianwouters [at] gmail.com

[1] L. D. Landau, Zh. Eksp. Teor. Fiz. **7**, 19 (1937); **7**, 627 (1937).

[2] D. R. Hartree, Math. Proc. Camb. Phil. Soc. **24**, 89 (1928); Math. Proc. Camb. Phil. Soc. **24**, 111 (1928); J. C. Slater, Phys. Rev. **32**, 339 (1928); V. Fock, Z. Phys. A **61**, 126 (1930).

[3] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[4] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory*, 1st ed. (Wiley New-York, 2000); T. Helgaker, P. Jørgensen, J. Olsen, and M. A. Ratner, Physics Today **54**, 52 (2001).

[5] I. Tamm, J. Phys. (USSR) **9**, 449 (1945); S. M. Dancoff, Phys. Rev. **78**, 382 (1950).

[6] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).

[7] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[8] E. U. Condon, Phys. Rev. **36**, 1121 (1930).

[9] F. Coester, Nucl. Phys. **7**, 421 (1958); F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960); J. Cizek, J. Chem. Phys. **45**, 4256 (1966).

[10] P. A. M. Dirac, Proc. R. Soc. Lond. A **114**, 243 (1927).

[11] V. Fock, Z. Phys. A **75**, 622 (1932).

[12] W. Dickhoff and D. Van Neck, *Many-Body Theory Exposed!*, 2nd ed. (World Scientific, 2008).

[13] D. J. Rowe, A. Ryman, and G. Rosensteel, Phys. Rev. A **22**, 2362 (1980); G. Rosensteel and D. J. Rowe, Phys. Rev. A **24**, 673 (1981).

[14] D. J. Thouless, Nucl. Phys. **21**, 225 (1960); Nucl. Phys. **22**, 78 (1961).

[15] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); Phys. Rev. B **48**, 10345 (1993); S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995); S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997); H. Saberi, A. Weichselbaum, and J. von Delft, Phys. Rev. B **78**, 035124 (2008);

M. B. Hastings, J. Stat. Mech. Theor. Exp. **2007**, P08024 (2007).

[16] X. G. Wen, Int. J. Mod. Phys. B **4**, 239 (1990); X. Chen, Z.-C. Gu, and X.-G. Wen, Phys. Rev. B **83**, 035107 (2011); L. Fidkowski and A. Kitaev, Phys. Rev. B **83**, 075103 (2011); A. M. Turner, F. Pollmann, and E. Berg, Phys. Rev. B **83**, 075102 (2011); N. Schuch, D. Pérez-García, and I. Cirac, Phys. Rev. B **84**, 165139 (2011); M. P. Zaletel and R. S. K. Mong, Phys. Rev. B **86**, 245305 (2012).

[17] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1999); G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).

[18] S. Wouters, P. A. Limacher, D. Van Neck, and P. W. Ayers, J. Chem. Phys. **136**, 134110 (2012); S. Sharma and G. K.-L. Chan, J. Chem. Phys. **136**, 124121 (2012).

[19] G. K.-L. Chan and S. Sharma, Annu. Rev. Phys. Chem. **62**, 465 (2011).

[20] G. K.-L. Chan, Phys. Chem. Chem. Phys. **10**, 3454 (2008).

[21] J. M. Kinder, C. C. Ralph, and G. K.-L. Chan, ArXiv e-prints (2011), arXiv:1103.2155 [cond-mat.str-el]; in *APS Meeting Abstracts* (2013) p. BAPS.2013.MAR.J24.6.

[22] P. A. M. Dirac, Math. Proc. Camb. Phil. Soc. **26**, 376 (1930); P. W. Langhoff, S. T. Epstein, and M. Karplus, Rev. Mod. Phys. **44**, 602 (1972); A. K. Kerman and S. E. Koonin, Ann. Phys. **100**, 332 (1976); P. Kramer and M. Saraceno, in *Group Theoretical Methods in Physics*, Lecture Notes in Physics, Vol. 135 (Springer-Verlag Berlin Heidelberg, 1980) pp. 112–121.

[23] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Phys. Rev. Lett. **107**, 070601 (2011).

[24] K. Ueda, C. Jin, N. Shibata, Y. Hieida, and T. Nishino, ArXiv e-prints (2006), arXiv:cond-mat/0612480 [cond-mat.stat-mech].

[25] J. Haegeman, *Variational renormalization group methods for extended quantum systems*, Ph.D. thesis, Ghent University (2011).

[26] P. S. Zyrianov and E. M. Eleonski, Soviet Phys. JETP **3**, 620 (1956); R. A. Ferrell, Phys. Rev. **107**, 450 (1957); Phys. Rev. **107**, 1631 (1957).

[27] K. Sawada, Phys. Rev. **106**, 372 (1957); K. Sawada, K. A. Brueckner, N. Fukuda, and R. Brout, Phys. Rev. **108**, 507 (1957); P. W. Anderson, Phys. Rev. **112**, 1900 (1958); H. Ehrenreich and M. H. Cohen, Phys. Rev. **115**, 786 (1959); J. Goldstone and K. Gottfried, Nuovo Cim. **13**, 849 (1959).

[28] J. J. Dorando, J. Hachmann, and G. K.-L. Chan, J. Chem. Phys. **130**, 184111 (2009).

[29] P. Ring and P. Schuck, *The nuclear many-body problem*, 1st ed. (Springer-Verlag Berlin Heidelberg, 1980).

[30] L. Brillouin, *La méthode du champ self-consistent*, Actualités Scientifiques et Industrielles, Vol. 71 (Hermann, 1933) Sections 159-160.

[31] Y. Nambu, Phys. Rev. **117**, 648 (1960); J. Goldstone, Nuovo Cim. **19**, 154 (1961).

[32] J. Haegeman, M. Mariën, T. J. Osborne, and F. Verstraete, ArXiv e-prints (2012), arXiv:1210.7710 [quant-ph].

[33] S. Holtz, T. Rohwedder, and R. Schneider, Numer. Math. **120**, 701 (2012); A. Uschmajew and B. Vandereycken, Linear Algebra Appl. **439**, 133 (2013).

[34] U. Schollwöck, Ann. Phys. **326**, 96 (2011).

[35] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

[36] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[37] J. Hubbard, Proc. R. Soc. Lond. A **276**, 238 (1963).

[38] B. Pirvu, J. Haegeman, and F. Verstraete, Phys. Rev. B **85**, 035130 (2012); J. Haegeman, B. Pirvu, D. J. Weir, J. I. Cirac, T. J. Osborne, H. Verschelde, and F. Verstraete, Phys. Rev. B **85**, 100408 (2012).

[39] I. P. McCulloch and M. Gulácsi, Europhys. Lett. **57**, 852 (2002); I. P. McCulloch, J. Stat. Mech.: Theory Exp. **10**, 14 (2007); S. Singh, R. N. C. Pfeifer, and G. Vidal, Phys. Rev. A **82**, 050301 (2010); S. Singh, H.-Q. Zhou, and G. Vidal, New J. Phys. **12**, 033029 (2010).

[40] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).

[41] C. N. Yang and S. C. Zhang, Mod. Phys. Lett. B **04**, 759 (1990).

[42] K. Ohno, Theor. Chim. Acta **2**, 219 (1964).

[43] I. Ohmine, M. Karplus, and K. Schulten, J. Chem. Phys. **68**, 2298 (1978).

[44] S. Ramasesha, S. K. Pati, H. R. Krishnamurthy, Z. Shuai, and J. L. Brédas, Phys. Rev. B **54**, 7598 (1996); Z. Shuai, S. K. Pati, W. P. Su, J. L. Brédas, and S. Ramasesha, Phys. Rev. B **55**, 15368 (1997); Z. Shuai, J. L. Brédas, S. K. Pati, and S. Ramasesha, Phys. Rev. B **56**, 9298 (1997); Y. Anusooya, S. K. Pati, and S. Ramasesha, J. Chem. Phys. **106**, 10230 (1997); C. Raghu, Y. Anusooya Pati, and S. Ramasesha, Phys. Rev. B **66**, 035116 (2002); S. Mukhopadhyay and S. Ramasesha, J. Chem. Phys. **131**, 074111 (2009).

[45] J. J. Dorando, J. Hachmann, and G. K.-L. Chan, J. Chem. Phys. **127**, 084109 (2007).

[46] P. Tavan and K. Schulten, Phys. Rev. B **36**, 4337 (1987).

[47] J. Haegeman, T. J. Osborne, and F. Verstraete, ArXiv e-prints (2013), arXiv:1305.1894.