UNIVERSITY OF CALIFORNIA,
IRVINE

Improving matrix product state methods for long-range interactions

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Mingru Yang

Dissertation Committee:
Distinguished Porfessor Steven R. White, Chair
Professor Sasha Chernyshev
Professor Kieron Burke

2020

DEDICATION

To

my grandparents

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor, Prof. Steven R. White, who shares many physical insights, gives me a freedom to explore my research interest, and treat me as an independent researcher.

I would like to thank my collaborators, Sean R. Muleady and Prof. Ana-Maria Rey at University of Colorado, Boulder, for teaching me quantum metrology.

I would also like to thank Prof. Sid Parameswaran, who encouraged me a lot at the beginning of my PhD years, as well as Prof. Sasha Chernyshev, Yuri Shirman, Arvind Rajaraman, Michael Ratz, Gary A. Chanan from Department of Physics and Astronomy, and Prof. Li-Sheng Tseng from Department of mathematics, for their excellent lectures. I need also thank Prof. Xenophon Zotos from University of Crete in Greece for teaching me Bethe ansatz.

The growth of a PhD usually involves the help from postdocs. In particular, I want to thank Edwin M. Stoudenmire and Matthew Fishman for various discussions on ITensor.

I need to thank my Japanese instructors Eiko H. Sithi-Amnuai and Ryuko A. Flores, who created a lot of fun in my life outside the research.

I also need to thank my friends in UCI, especially Pavel Maksimov for helping to organize the condensed matter journal club, as well as Brian Casas and Randy Sawaya for helping me improve my oral English.

Above all, I am grateful to my parents for supporting my interest, respecting my personality, and their efforts in providing me with better education.

The work in this dissertation is supported by National Science Foundation through grant No. DMR-1812558 and the many electron collaboration of Simons Foundation .

VITA

Mingru Yang

EDUCATION

Doctor of Philosophy in Physics **2020**
University of California, Irvine *Irvine, California*

Bachelor of Science in Physics **2015**
Peking University *Beijing, China*

Minor in Computer Science **2013-2015**
Peking University *Beijing, China*

RESEARCH EXPERIENCE

Graduate Student Researcher **2015–2020**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2018**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

Time-dependent variational principle with ancillary Krylov subspace **2020**
Physical Review B 102, 094315

Density-matrix-renormalization-group study of a one dimensional diatomic molecule beyond the Born-Oppenheimer approximation **2019**
Physical Review A 99, 022509

SOFTWARE CONTRIBUTIONS

ITensor/ITensor `https://github.com/ITensor/ITensor`
A C++ library for implementing tensor network calculations

ITensor/TDVP `https://github.com/ITensor/TDVP`
A C++ library for implementing time-dependent variational principle for finite matrix product states

ABSTRACT OF THE DISSERTATION

Improving matrix product state methods for long-range interactions

By

Mingru Yang

Doctor of Philosophy in Physics

University of California, Irvine, 2020

Distinguished Porfessor Steven R. White, Chair

We study one dimensional models of diatomic molecules where both the electrons and nuclei are treated as quantum particles, going beyond the usual Born-Oppenheimer approximation. The continuous system is approximated by a grid which computationally resembles a ladder, with the electrons living on one leg and the nuclei on the other. To simulate DMRG efficiently with this system, a three-site algorithm has been implemented. We also use a compression method to treat the long-range interactions between charged particles. We find that 1D diatomic molecules with spin-1/2 nuclei in the spin-triplet state will unbind when the mass of the nuclei reduces to only a few times larger than the electron mass, while the molecule with nuclei in the singlet state always binds, given the two electrons in their singlet state in both cases.

We propose an improved scheme to do the time dependent variational principle (TDVP) in finite matrix product states (MPS) for two-dimensional systems or one-dimensional systems with long range interactions. We present a method to represent the time-evolving state in a MPS with its basis enriched by state-averaging with global Krylov vectors. We show that the projection error is significantly reduced so that precise time evolution can still be obtained even if a larger time step is used. Combined with the one-site TDVP, our approach provides a way to dynamically increase the bond dimension while still preserving unitarity for real

time evolution. Our method can be more accurate and exhibit slower bond dimension growth than the conventional two-site TDVP.

We apply our improved TDVP method to investigate the spin squeezing dynamics of the XXZ model of $1/r^{\alpha}$ interaction in two dimension. Comparing with the spin squeezing parameter and other observables obtained from discrete truncated Wigner approximation (DTWA), we verify the validity of DTWA and unveil the potential for this method to study dynamics of large-scale spin systems. Our results confirm the existence of large collective regime when $\alpha > 2$, which can be a guide for future experimental realizations. Combined with the purification method, our improved TDVP method is proved to be also useful to study the thermalization of the long-range interacting system.

Chapter 1

Preliminaries

1.1 Introduction

The field of tensor networks was started from the invention of the density matrix renormalization group (DMRG) method. In the paper[1] by White in 1992, DMRG was proved to be effective to solve the ground state of Heisenberg chains, which overcame the drawbacks of the Wilson's numerical renormalization group[2, 3] in this problem. After various numerical experiments of DMRG on one-dimensional and two-dimensional systems, it was found that while DMRG appears to be perfect for one-dimensional gapped ground states, it is only partially successful for two-dimensional systems and critical systems. Borrowing concepts from quantum information, people realized[4, 5, 6, 7] that the efficiency of DMRG on one dimensional gapped ground states relies on the constant entanglement entropy which does not scale with system size in one dimension according to the entanglement area law[8, 9].

In the past two decades, there has been tremendous development of tensor network theories and algorithms. In the late 90's, people began to reformulate[10, 11] DMRG in terms of matrix product states (MPS), which actually have a long history in analytical calculations[12,

13, 14], although under different names. Rethinking in the MPS language boosted the inventions of various new extensions of DMRG at the beginning of this century. For example, the time evolution methods[15, 16, 17, 18], infinite-system algorithms[19, 20, 21, 22], continuous systems[23], and higher-dimensional tensor networks such as projected entangled pair states (PEPS)[24, 25, 26] and multi-scale entanglement renormalization ansatz (MERA)[27] as well as continuous extensions of them[28, 29].

Tensor networks not only serve as a mathematical representation of quantum states, but also have been proved to be useful in exploring the physics in many problems by numerical simulations. For example, it has been used to study the properties of the ground state of Hubbard model to understand high-T_c superconductivity[30]. It has been also applied to searching for topological ordered states like spin liquids[31] and fractional quantum Hall ground states[32]. Along with the development of compact basis functions[33] and effective compression of matrix product operators, it further shows the ability to tackle quantum chemistry problems[34, 35]. In addition, tensor networks interact actively with machine learning[36, 37] and quantum computation[38].

Long in the history of quantum chemistry and condensed matter physics, the nuclear degrees of freedom of atoms has been separated out from the electronic states by the Born-Oppenheimer approximation. However, the non-adiabatic effects due to nuclei has been proved to be important[39, 40] in chemical reactions when conical intersections occur in potential energy surfaces[41] and exotic molecules with the mass of the nucleus close to that of the electron. We thus utilized the matrix product states techniques to study the ground states of the system with two types of fermions, including the nuclear degrees of freedom by full quantum mechanical treatment[42] and discovered an unconventional binding mechanism. In this work, we extend the conventional two-site DMRG to three-site in order to handle this special system which has no hopping between the nuclei and electrons, hence avoiding the local-minimum problem. When the mass ratio between the nuclei and electrons

is large, there is a problem of slow convergence in DMRG due to the existence of two distinct energy scales. We came up with the idea to use different lattice spacing for each species of particle, which solves the problem after combining it with swap gates. Similar pitfall also exists in time-dependent variational princple, especially for long-range interactions and two-dimensional systems. Not like ground-state calculations, where only the final convergence to the ground state energy is important, in the time evolution, every time step (sweep) need to be accurate since we are interested in the whole dynamical process. Then we would ask if the tricks in DMRG could also be used in TDVP to solve the problem. It is not clear immediately how to integrate the swap gates into the derivation of effective TDVP equations, and extension to more center sites will increase the complexity. Alternatively, inspired by the multi-state targeting in DMRG[43], we found that subspace expansion combined with single-site TDVP appears to be quite efficient and reliable[44] and successfully applied it to study the spin squeezing dynamics of long-range XXZ model in two dimension.

In this chapter, we will briefly review the MPS techniques for ground state calculation and time evolution, as preliminary to the chapters followed. For more details, we recommend the reader to refer to Ref. [45].

1.2 Entanglement entropy

For a bipartition of a quantum system into A and B, the von Neumann entropy is defined as

$$S_{A|B} = -\mathrm{Tr}\rho_A \log_2 \rho_A, \tag{1.1}$$

where ρ_A is the reduced density matrix of A by tracing out the subsystem B, i.e.

$$\rho_A = \mathrm{Tr}_B \rho, \tag{1.2}$$

where ρ is the density matrix of the whole system. For a pure state $|\psi\rangle$, we have $\rho = |\psi\rangle\langle\psi|$. If we write the pure state as

$$|\psi\rangle = \sum_{i=1}^{D_A} \sum_{j=1}^{D_B} \Psi_{ij} |i\rangle_A |j\rangle_B, \tag{1.3}$$

where $|i\rangle_A$ and $|j\rangle_B$ form orthonormal bases of A and B respectively, then ρ_A can be represented as a positive-semidefinite matrix in the subsystem basis $|i\rangle_A$ and $|j\rangle_B$

$$\rho_A = \Psi \Psi^\dagger. \tag{1.4}$$

We can diagonalize the reduced density matrix ρ_A

$$\rho_A = U \mathrm{diag}(w_1, w_2, ...) U^\dagger = \sum_{a=1}^{r} w_a |a\rangle_{AA} \langle a|, \tag{1.5}$$

where $U = (|1\rangle_A, |2\rangle_A, ...)$ and $r = \min(D_A, D_B)$. If the whole system is in a pure state, equivalently we can obtain the Schmidt decomposition

$$|\psi\rangle = \sum_{a=1}^{r} s_a |a\rangle_A |a\rangle_B, \tag{1.6}$$

where $|a\rangle_A = \sum_i U_{ia} |i\rangle_A$ and $|a\rangle_B = \sum_j V^*_{ja} |j\rangle_A$, by performing a singular value decomposition (SVD) to Ψ, i.e.

$$\Psi = U S V^\dagger, \tag{1.7}$$

where $S = \mathrm{diag}(s_1, s_2, ...) = \mathrm{diag}(\sqrt{w_1}, \sqrt{w_2}, ...)$; U has orthonormal columns (left orthonormal), i.e. $U^\dagger U = I$; $V^\dagger$ has orthonormal rows (right orthonormal), i.e. $V^\dagger V = I$. If $D_A > D_B$, $VV^\dagger = I$ and $UU^\dagger = P_L$, where P_L is the projection operator into the column space of Ψ. If $D_A < D_B$, $UU^\dagger = I$ and $VV^\dagger = P_R$, where P_R is the projection operator into the row space of Ψ. Therefore the von Neumann entropy is directly represented by the eigenvalues of the reduced density matrix or the singular values from SVD of the pure state, i.e.

$$S_{A|B} = -\sum_a w_a \log_2 w_a. \tag{1.8}$$

If the system is in pure state, the von Neumann entropy[1] characterizes the entanglement between the subsystem A and B. For a mixed state, the von Neumann entropy contains a thermal contribution in addition to the entanglement. While the thermal entropy is extensive, it is not necessarily true for the entanglement entropy. In the thermodynamic limit, the scaling of the entanglement entropy of a ground state obeys the area law[7, 8, 9] if the ground state is gapped out from excitations for a short-ranged Hamiltonian. So $S_{A|B} \sim$ const in 1D and $S_{A|B} \sim L$ in 2D, where L is the length of the boundary between A and B. At critical point, $S_{A|B}$ has a logarithmic correction $\frac{c+\bar{c}}{6} \log_2 L$ in 1D from conformal field theory[4] analysis, and distinct behaviors are shown[46] for bosonic[47] and fermionic[48] systems in higher dimensions.

If the local Hilbert space of a quantum lattice system is d dimensional, generally we need d^r Schmidt states to represent a pure state $|\psi\rangle$ exactly, which scales exponentially with the system size. For the ground states, however, the speciality of area laws make it possible to efficiently and accurately approximate them in computationally reachable finite number of Schmidt states, which is the key to the success of the density matrix renormalization group

[1]The second Renyi entropy is another choice which is more popular in analytical calculations. The von Neumann entropy is more sensitive to small tails of singular values

(DMRG) and matrix product states (MPS).

If we use $D \leq r$ Schmidt states to approximate a pure state $|\psi\rangle$, the maximal entanglement it can encode is $S_{A|B} = \log_2 D$ when all the eigenvalues of the reduced density matrix ρ_A are equal. This implies we need to use at least $D = 2^{\text{const}}$ ($D = 2^L$) number of Schmidt states to encode a 1D (2D) gapped ground state, and $D = L^{\frac{c+\bar{c}}{6}}$ for a 1D critical system. Increase in 1D gapped system size does not affect the lower bound, while the infinite lower bound when $L \to \infty$ means the thermodynamic limit is not reachable for 1D critical system. Increase in the boundary length between partitions of 2D gapped systems will lead to an exponential increase in D. Although the real distribution of eigenvalues of ρ_A are generally far from uniform[49], the lower bound D still provides a good estimate of the necessary number of Schmidt states, which is verified by the fact that the eigenvalues w_a decay exponentially for gapped chains and stripes with an increasing decay length as the width of the system grows.

1.3 Matrix product states

A many-body state with open boundary conditions for a lattice of N number of sites with the local physical degrees of freedom labeled by s_n is given by

$$|\psi\rangle = \sum_{s_1 \ldots s_N} c_{s_1 \ldots s_N} |s_1 \ldots s_N\rangle, \tag{1.9}$$

which can be decomposed as an finite matrix product state (MPS) as

$$|\psi[M]\rangle = \sum_{s_1 \ldots s_N} M_1^{s_1} \cdots M_N^{s_N} |s_1 \ldots s_N\rangle, \tag{1.10}$$

where M_n ($\forall n \in \{1, \ldots, N\}$) are rank-3 tensors. An entry of it can be written as $[M_n]^{s_n}_{b_{n-1} b_n}$, where s_n is the physical index of dimension d_n and b_{n-1} and b_n are the left and right virtual

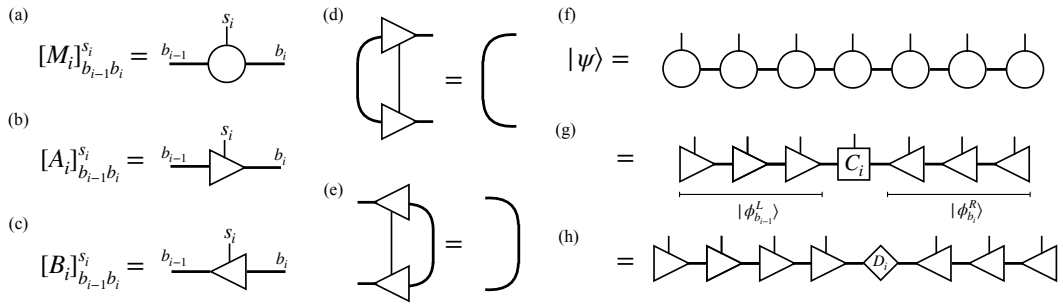

Figure 1.1: Graphic notations for an MPS.

indices of dimension m_{n-1} and m_n respectively. Usually $d_n \equiv d$ but m_n varies with n. For finite MPS, $m_0 = m_{N+1} = 1$ and $\max\{m_n\}$ is reached in the middle of the lattice. m_n is conventionally called the bond dimension for the nth bond of the MPS.

The MPS representation has a redundancy called gauge freedom, i.e. the physical state $|\psi\rangle$ is invariant under the gauge transformation

$$M_n^{s_n} \mapsto M_n'^{s_n} = G_n^L M_n^{s_n} G_n^R, \tag{1.11}$$

where $G_{n-1}^R G_n^L = I$ and $G_1^L = G_N^R = I$. We can bring the MPS to canonical forms by fixing the gauge. For example, to get the mixed canonical form with the orthogonality center at site i

$$|\psi\rangle = \sum_{s_1 \ldots s_N} A_1^{s_1} \cdots A_{i-1}^{s_{i-1}} C_i^{s_i} B_{i+1}^{s_{i+1}} \cdots B_N^{s_N} |s_1 \ldots s_N\rangle, \tag{1.12}$$

where $A_n^{s_n}$ is left-orthonormal and $B_n^{s_n}$ is right-orthonormal, i.e.

$$\begin{aligned} \sum_{s_n} (A_n^\dagger)^{s_n} A_n^{s_n} &= I, \\ \sum_{s_n} B_n^{s_n} (B_n^\dagger)^{s_n} &= I, \end{aligned} \tag{1.13}$$

we can fix the gauge by the singular value decomposition (SVD) of the site tensors first from

$n = 1$ to $n = N$ iteratively, i.e.

$$A_n^{s_n} G_{n+1}^L = U_n S_n V_n = G_n^L M_n^{s_n}, \tag{1.14}$$

where $G_n^L = S_{n-1} V_{n-1}$ to get the left-canonical form, and then SVD from $n = N$ to $n = i$ iteratively, i.e.

$$G_{n-1}^R B_n^{s_n} = U_n S_n V_n = A_n^{s_n} G_n^R, \tag{1.15}$$

where $G_n^R = U_{n+1} S_{n+1}$ and we have

$$C_i^{s_i} = A_i^{s_i} D_i = D_{i-1} B_i^{s_i}, \tag{1.16}$$

where $D_i = U_{i+1} S_{i+1}$. Eq (1.12) can also be rewritten as

$$|\psi\rangle = \sum_{b_{i-1} s_i b_i} [C_i]^{s_i}_{b_{i-1} b_i} |\phi^L_{b_{i-1}}\rangle |s_i\rangle |\phi^R_{b_i}\rangle, \tag{1.17}$$

where

$$\begin{aligned} |\phi^L_{b_{i-1}}\rangle &= \sum_{s_1 \ldots s_{i-1}} (A_1^{s_1} \ldots A_{i-1}^{s_{i-1}})_{b_{i-1}} |s_1 \ldots s_{i-1}\rangle \\ &= \sum_{b_{i-2}, s_{i-1}} \sum_{s_1 \ldots s_{i-2}} (A_1^{s_1} \ldots A_{i-2}^{s_{i-2}})_{b_{i-2}} |s_1 \ldots s_{i-2}\rangle (A_{i-1}^{s_{i-1}})_{b_{i-2}, b_{i-1}} |s_{i-1}\rangle \\ &= \sum_{b_{i-2}, s_{i-1}} (A_{i-1}^{s_{i-1}})_{b_{i-2}, b_{i-1}} |\phi^L_{b_{i-2}}\rangle |s_{i-1}\rangle \end{aligned} \tag{1.18}$$

and similarly

$$\begin{aligned} |\phi^R_{b_i}\rangle &= \sum_{s_{i+1} \ldots s_N} (B_{i+1}^{s_{i+1}} \ldots B_N^{s_N})_{b_i} |s_{i+1} \ldots s_N\rangle \\ &= \sum_{s_{i+1}, b_{i+1}} (B_{i+1}^{s_{i+1}})_{b_i, b_{i+1}} |s_{i+1}\rangle |\phi^R_{b_{i+1}}\rangle \end{aligned} \tag{1.19}$$

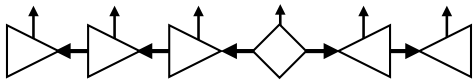

Figure 1.2: Quantum number currents in the canonical form of a MPS.

are automatically orthonormal bases for the left and right partition of the lattice respectively.

For later discussions, we introduce here how a global Abelian symmetry is implemented in the MPS formulation. Let us consider the magnetization $U(1)$ symmetry as an example. Suppose that the local basis $|s_i\rangle$ at each site is chosen as a eigenstate of $\hat{S}_i^z$ with eigenvalue M_{s_i}, and the left (right) bond basis $|\phi^L_{b_{i-1}}\rangle$ ($|\phi^R_{b_i}\rangle$) is chosen as a eigenstate of $\sum_{i=1}^{i-1}\hat{S}_j^z$ ($\sum_{i=i+1}^{N}\hat{S}_j^z$) with eigenvalue $M^L_{b_{i-1}}$ ($M^R_{b_i}$). We then have

$$\begin{aligned}\sum_{i=1}^{i-1}\hat{S}_j^z|\phi^L_{b_{i-1}}\rangle = M^L_{b_{i-1}}|\phi^L_{b_{i-1}}\rangle &= \sum_{b_{i-2},s_{i-1}} M^L_{b_{i-1}}(A_{i-1}^{s_{i-1}})_{b_{i-2},b_{i-1}}|\phi^L_{b_{i-2}}\rangle|s_{i-1}\rangle \\ &= (\sum_{i=1}^{i-2}\hat{S}_j^z + \hat{S}_i^z)\sum_{b_{i-2},s_{i-1}}(A_{i-1}^{s_{i-1}})_{b_{i-2},b_{i-1}}|\phi^L_{b_{i-2}}\rangle|s_{i-1}\rangle \\ &= \sum_{b_{i-2},s_{i-1}}(M^L_{b_{i-2}}+M_{s_{i-1}})(A_{i-1}^{s_{i-1}})_{b_{i-2},b_{i-1}}|\phi^L_{b_{i-2}}\rangle|s_{i-1}\rangle,\end{aligned} \tag{1.20}$$

which actually imposes a condition for the non-vanishing entries of $(A_{i-1}^{s_{i-1}})_{b_{i-2},b_{i-1}}$, i.e.

$$M^L_{b_{i-1}} = M^L_{b_{i-2}} + M_{s_{i-1}}. \tag{1.21}$$

Similarly we have

$$M^R_{b_i} = M^R_{b_{i+1}} + M_{s_{i+1}}. \tag{1.22}$$

It implies that the matrices $[A_{i-1}]_{b_{i-2}s_{i-1},b_{i-1}}$ and $[B_{i+1}]_{b_i,s_{i+1}b_{i+1}}$ take a block structure. The above equations can be regarded as the continuity equation if we think of the quantum

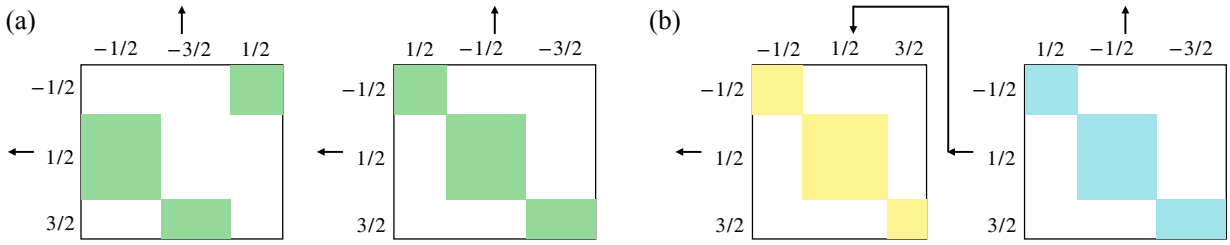

Figure 1.3: (a) An example of the quantum number blocks of the orthogonality center of a MPS that has $M = 0$. After rearrangement of the column and rows, the orthogonality center becomes block diagonal. (b) The quantum number structure is unchanged after a SVD.

numbers as "curents" (see FIG. 1.2). If $|\psi\rangle$ has a total magnetization M, we will have

$$\begin{aligned}\sum_{j=1}^{N} \hat{S}_j^z |\psi\rangle &= M|\psi\rangle \\ &= (\sum_{i=1}^{i-1} \hat{S}_j^z + \hat{S}_i^z + \sum_{i=i+1}^{N} \hat{S}_j^z) \sum_{b_{i-1} s_i b_i} [C_i]^{s_i}_{b_{i-1} b_i} |\phi^L_{b_{i-1}}\rangle |s_i\rangle |\phi^R_{b_i}\rangle \\ &= \sum_{b_{i-1} s_i b_i} (M^L_{b_{i-1}} + M_{s_i} + M^R_{b_i}) [C_i]^{s_i}_{b_{i-1} b_i} |\phi^L_{b_{i-1}}\rangle |s_i\rangle |\phi^R_{b_i}\rangle,\end{aligned} \tag{1.23}$$

i.e. for each non-zero element $[C_i]^{s_i}_{b_{i-1} b_i}$, we need

$$M = M^L_{b_{i-1}} + M_{s_i} + M^R_{b_i}, \tag{1.24}$$

which means the orthogonality center is the "source" of the quantum number "currents".

The conditions Eq. (1.21), (1.22), and (1.24) guarantees that the site tensors can be taken to be block diagonal (FIG. 1.3(a)). Thus a SVD of C_i is also in fact carried out block-diagonally, thus retaining the quantum number structure (FIG. 1.3(b)).

1.4 MPS compression methods

There are three ways[45, 50] to compress a MPS of bond dimension D to D'. They are SVD compression[45], density-matrix compression[51], and variational compression[52, 45]. Those compression methods are not limited to the situation of having two MPSs, but can be adapted slightly in e.g. MPO application[50, 53] to an MPS and summation of MPSs[45].

The SVD compression is suitable for the case when the compression is slight and the MPS has already in its canonical form. It has the disadvantage of one-sided dependence which becomes especially serious if the compression is strong. The typical complexity is $O(dD'D^2)$

The density-matrix compression is equivalent to SVD when the MPS is canonical, but it can also work for a non-canonical MPS, for example in the MPO-MPS application. Similar issue of one-sided dependence happens in the density-matrix compression if the compression is strong[2]. The time and space complexity of this method is the highest ($O(dD^3)$), due to the construction of the edge tensors.

The variational way works similar as DMRG. It variationally optimize the 2-norm distance $\||\psi\rangle - |\tilde{\psi}\rangle\|^2$. When the MPS is kept in the canonical form, we do not need to solve a linear equation. Although generally the time complexity for the matrix-vector multiplication is still $O(dD'D^2)$, in specific cases like MPS summation[45] and MPO-MPS application[50], there are tricks to reduce the complexity by utilizing their special structure. Just like DMRG, the variational compression has a pitfall to get stuck in the local minimum and its convergence speed largely depends on the initial guess. One can use the two-site variant but the local minimum problem is not guaranteed to be solved for some systems. Usually we can first do a SVD compression to get a good ansatz and use it as the input for the variational compression.

[2]In our GSE-TDVP calculation of the 2D long-range XXZ model, we actually found the side-dependence of $\langle \hat{S}_i^x \rangle$ for $\alpha > 3$ near the Heisenberg point, which indicates that a smaller truncation should be used in the global subspace expansion.

We also noticed recently there is a tangent space method to compress the uniform MPS[54].

1.5 Matrix product operators

There is also a similar MPO representation for an operator

$$\hat{O} = \sum_{s_1 \ldots s_N, s'_1 \ldots s'_N} W_1^{s_1 s'_1} \cdots W_N^{s_N s'_N} |s_1 \ldots s_N\rangle\langle s'_1 \ldots s'_N|. \tag{1.25}$$

Especially we consider the MPO representation of a Hamiltonian. Any Hamiltonian for a 1D finite lattice system can be decomposed[17] as

$$\hat{H} = \hat{H}_{L_i} \otimes \hat{1}_{R_i} + \hat{1}_{L_i} \otimes \hat{H}_{R_i} + \sum_{a_i=1}^{n_i} \hat{h}_{L_i,a_i} \otimes \hat{h}_{R_i,a_i}, \tag{1.26}$$

where L_i (R_i) is the part of the system left (right) to the bond between site i and $i+1$. The recursion relation

$$\begin{pmatrix} \hat{H}_{R_{i-1}} \\ \hat{h}_{R_{i-1}} \\ \hat{1}_{R_{i-1}} \end{pmatrix} = \begin{pmatrix} \hat{1}_i & \hat{C}_i & \hat{D}_i \\ 0 & \hat{A}_i & \hat{B}_i \\ 0 & 0 & \hat{1}_i \end{pmatrix} \otimes \begin{pmatrix} \hat{H}_{R_i} \\ \hat{h}_{R_i} \\ \hat{1}_{R_i} \end{pmatrix} \tag{1.27}$$

can be used to generate the site tensor W_i of the MPO of the Hamiltonian. The generation of the MPO through the recursion relation (1.27) can be interpreted as a finite state automaton[55, 56, 45], as illustrated in FIG. 1.4. The MPO constructed in this way will have a bond dimension $D_W = n_i + 2$.

The bond dimension D_W of an MPO can grow rapidly with the interaction range. However,

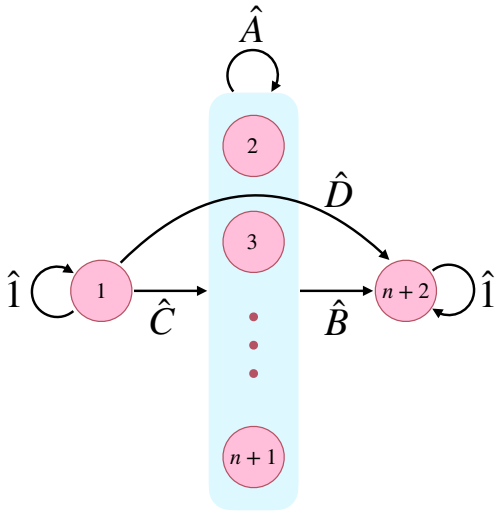

Figure 1.4: Finite state automaton to construct the MPO of the Hamiltonian $\hat{H}$. The nodes denote the $n_i + 2$ bond indices of W_i, and the directed edges are the matrix operator entries of W_i. Each path started from node 1 and ended at node $n + 2$ gives a string of matrix operators at each site. The summation of all the paths generates the Hamiltonian.

there are exceptions. For example, if $\hat{H} = \sum_{ij} \hat{S}_i^z \hat{S}_j^z$, we will have

$$W_i = \begin{pmatrix} \hat{1}_i & 2\hat{S}_i^z & (\hat{S}_i^z)^2 \\ 0 & \hat{1}_i & \hat{S}_i^z \\ 0 & 0 & \hat{1}_i \end{pmatrix}, \tag{1.28}$$

i.e. the bond dimension $D_W = 3$ although the interaction is infinite-range. Another important exception is the exponentially decaying interaction. Consider $\hat{H} = \sum_{ij} e^{-|i-j|/\xi} \hat{S}_i^z \hat{S}_j^z$, and we will have an MPO of $D_W = 3$, i.e.

$$W_i = \begin{pmatrix} \hat{1}_i & 2\lambda\hat{S}_i^z & (\hat{S}_i^z)^2 \\ 0 & \lambda\hat{1}_i & \hat{S}_i^z \\ 0 & 0 & \hat{1}_i \end{pmatrix}, \tag{1.29}$$

where $\lambda = e^{-1/\xi}$. More generally, we need to compress the MPO for long-range interactions for computation efficiency. One option[55, 57] is approximating the interaction strength $J(r)$ by a sum of exponentials and minimizing the 2-norm distance $\|J(r) - \sum_{i=1}^{m} \alpha_i \lambda_i^r\|$, where

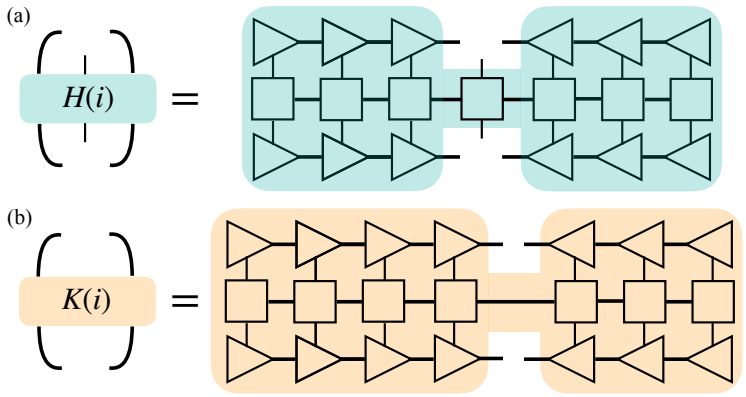

Figure 1.5: Graphic notations for the single-site and zero-site effective Hamiltonians.

m depends on the accuracy required and the targeting D_W. Another option[58] would be adapting the MPS compression methods to reduce the bond dimension of the MPO.

We can also project the Hamiltonian to the space spanned by the bond basis (orthonormal Schmidt states) and get the effective Hamiltonians (as shown in FIG. 1.5). In terms of the canonical form of the MPS, the single-site effective Hamiltonian $H(i)$ can be written as

$$[H(i)]_{b'_{i-1}s'_ib'_i;b_{i-1}s_ib_i} = \langle\phi^R_{b'_i}|\langle s'_i|\langle\phi^L_{b'_{i-1}}|\hat{H}|\phi^L_{b_{i-1}}\rangle|s_i\rangle|\phi^R_{b_i}\rangle. \tag{1.30}$$

Similarly using Eq (1.16) we can rewrite

$$|\psi\rangle = \sum_{a_ib_i}[D_i]_{a_ib_i}|\phi^L_{a_i}\rangle|\phi^R_{b_i}\rangle, \tag{1.31}$$

and get the zero-site effective Hamiltonian $K(i)$

$$[K(i)]_{a'_ib'_i;a_ib_i} = \langle\phi^R_{b'_i}|\langle\phi^L_{a'_i}|\hat{H}|\phi^L_{a_i}\rangle|\phi^R_{b_i}\rangle. \tag{1.32}$$

1.6 Density matrix renormalization group

DMRG is a powerful algorithm to find the ground state of 1D systems and 2D stripes, the success of which is based on the representability of MPSs for states that obey the entanglement area law. Here we will only focus on the DMRG for finite MPSs. The goal is to variationally optimize an MPS $|\psi\rangle$ to minimize the energy

$$E = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}, \tag{1.33}$$

i.e. we need to minimize $\langle\psi|\hat{H}|\psi\rangle$ with the constraint $\langle\psi|\psi\rangle = 1$. We can transform it into an unconstrained optimization problem by introducing a Lagrange multiplier

$$\mathcal{L} = \langle\psi|\hat{H}|\psi\rangle - \lambda\langle\psi|\psi\rangle. \tag{1.34}$$

The single-site DMRG considers as variational parameters only one site tensor of the MPS at a step. If the MPS is not in its canonical form, $\partial\mathcal{L}/\partial M_i = 0$ will give us a generalized eigenvalue problem

$$H(i)M_i - \lambda N(i)M_i = 0, \tag{1.35}$$

where $H(i) = L_{i-1}W_iR_i$ and $N(i) = (M_{i-1}^\dagger \cdots M_1^\dagger M_1 \cdots M_{i-1})(M_{i+1} \cdots M_N M_N^\dagger \cdots M_{i+1}^\dagger)$. However, if we maintain the canonical form of the MPS which is centered at site i, we will have $N(i) = I$ so we only need to solve a standard eigenvalue problem

$$H(i)C_i - \lambda C_i = 0 \tag{1.36}$$

by iterative large sparse matrix eigensolvers like Lanczos or Jacobi-Davidson and get the smallest eigenvalue and the corresponding eigenvector. Then we can use the gauge transfor-

mation in Eq. (1.16) to move the orthogonality center to the next site, which not only keeps the canonical form at each step but also yields a valid starting vector at the next step. We repeat this procedure by sweeping back and forth in the system until convergence has been reached[3].

This single-site version can suffer seriously from the local minimum problem. As we have known, neither the SVD nor the Lanczos can increase the bond dimension of each quantum number sector or introduce new quantum number sectors to the subsystem. There are several ways to solve the problem. The simplest thing one can do is artificially enlarging the bond dimension by filling up zeros, but in practice it usually does not work well since it does not effectively introduce the necessary quantum number sectors. One can of course use the two-site variant of DMRG to allow bond dimension growth and new quantum number sectors[4] in the middle bond of the two center sites in the Lanczos and SVD, but at a cost of $O(d)$ times complexity.

Keeping in the one-site frame, there is a trick called the noise term[59] or equivalently the subspace expansion[60] in the MPS language, which can effectively help get out of the local minimum. The key observation is that the quantum number of the subsystem A is changed by quantum fluctuations due to the interaction between A and B. The MPO of the Hamiltonian can be decomposed as

$$\hat{H} = \sum_{a_i=1}^{D_W} \hat{H}^A_{a_i} \hat{H}^B_{a_i}, \tag{1.37}$$

[3]We can check $\langle\psi|\hat{H}^2|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle^2$ to see if the MPS has been a good eigenstate

[4]Whether this strategy works also depends on the Hamiltonian. It can introduce quantum number fluctuation for nearest-neighbor interactions, but there is a problem if the shortest-range interaction is the next-nearest-neighbor.

where

$$\begin{aligned}\hat{H}^A_{a_i} &= \sum_{s_1\ldots s_i, s'_1\ldots s'_i} (W_1^{s_1 s'_1}\cdots W_i^{s_i s'_i})_{a_i} |s_1\ldots s_i\rangle\langle s'_1\ldots s'_i|, \\ \hat{H}^B_{a_i} &= \sum_{s_{i+1}\ldots s_N, s'_{i+1}\ldots s'_N} (W_{i+1}^{s_{i+1} s'_{i+1}}\cdots W_N^{s_N s'_N})_{a_i} |s_{i+1}\ldots s_N\rangle\langle s'_{i+1}\ldots s'_N|.\end{aligned} \tag{1.38}$$

After a Lanczos optimization at site i, we can project the action of $\hat{H}^A_{a_i}$ on $|\psi\rangle$ to the space spanned by $\{|\phi^L_{b_{i-1}}\rangle|s_i\rangle|\phi^R_{b_i}\rangle\}$ and get

$$[P\hat{H}^A_{a_i}|\psi\rangle]_{b'_{i-1}s'_i,b_i} = [H^A_{a_i}]_{b'_{i-1}s'_i,b_{i-1}s_i}[C_i]^{s_i}_{b_{i-1}b_i}, \tag{1.39}$$

where $[H^A_{a_i}]_{b'_{i-1}s'_i,b_{i-1}s_i} = \langle s'_i|\langle\phi^L_{b'_{i-1}}|\hat{H}^A_{a_i}|\phi^L_{b_{i-1}}\rangle|s_i\rangle = L_{i-1}W_i$. We can form a density matrix from it

$$\tilde{\rho}_A = \sum_{a_i=1}^{D_W} \mathrm{Tr}_B\left[P\hat{H}^A_{a_i}|\psi\rangle\langle\psi|\hat{H}^A_{a_i}P\right]. \tag{1.40}$$

Then diagonalizing

$$\rho'_A = \rho_A + \alpha\tilde{\rho}_A = A_i S_i^2 A_i^\dagger \tag{1.41}$$

is equivalent to SVD

$$\begin{pmatrix} C_i & \sqrt{\alpha}L_{i-1}W_iC_i \end{pmatrix} = A_i S_i V_i, \tag{1.42}$$

which will enlarge the column space of C_i and thus introduce new basis to the ith bond if originally $b_{i-1}s_i < b_i$ and $H^A_{a_i}$ takes C_i out of the original column space. α is chosen to be a small number to make sure not too many weights coming from the original reduced density matrix $\rho_A = \mathrm{Tr}_B|\psi\rangle\langle\psi|$ are dropped out in the truncation. To move the orthogonality center

to the next site, we need

$$C_{i+1} = A_i^\dagger C_i B_{i+1}. \tag{1.43}$$

Notice that we cannot multiply $S_i V_i$ directly to B_{i+1} as in the ordinary way since their index dimensions do not match.

1.7 Time evolution methods

There are various methods to time evolve a MPS which can be classified into two classes. One class of methods like time-evolving block decimation (TEBD) or MPO $W^{\mathrm{I,II}}$ is trying to approximate the time evolution operator, while the other class of methods aim to solve the time-dependent Schrödinger equation globally or locally.

1.7.1 Time-evolving block decimation

TEBD can be implemented in the MPS language[16] and the classical system-environment DMRG language (tDMRG)[61]. The key idea is the Trotter-Suzuki decomposition of the time evolution operator $\hat{U}(\delta) = e^{-\mathrm{i}\delta\hat{H}}$ of a time step δ.

Suppose we have a Hamiltonian that consists nearest-neighbor interactions

$$\hat{H} = \sum_{i=1}^{N-1} \hat{h}_{i,i+1}, \tag{1.44}$$

which can be further decomposed as

$$\hat{H} = \hat{H}_{odd} + \hat{H}_{even} = \sum_{i\ odd} \hat{h}_{i,i+1} + \sum_{i\ even} \hat{h}_{i,i+1}. \tag{1.45}$$

To leading order, the Baker-Campbell-Hausdorff formula gives

$$\begin{aligned} e^{-\mathrm{i}\delta\hat{H}} &= e^{-\mathrm{i}\delta(\hat{H}_{odd}+\hat{H}_{even})} \approx e^{-\mathrm{i}\delta\hat{H}_{odd}} e^{-\mathrm{i}\delta\hat{H}_{even}} e^{\frac{\delta^2}{2}[\hat{H}_{odd},\hat{H}_{even}]} \\ &\approx e^{-\mathrm{i}\delta\hat{H}_{odd}} e^{-\mathrm{i}\delta\hat{H}_{even}} \left(1 + \frac{\delta^2}{2}[\hat{H}_{odd}, \hat{H}_{even}]\right) \\ &= e^{-\mathrm{i}\delta\hat{H}_{odd}} e^{-\mathrm{i}\delta\hat{H}_{even}} + O(\delta^2). \end{aligned} \tag{1.46}$$

Inside $\hat{H}_{odd}$, each term commutes with each other, so

$$e^{-\mathrm{i}\delta\hat{H}_{odd}} = \prod_{i \ odd} e^{-\mathrm{i}\delta\hat{h}_{i,i+1}}, \tag{1.47}$$

and similarly for $\hat{H}_{even}$. We actually obtain a first-order TEBD time stepper

$$\hat{U}^{\mathrm{TEBD1}}(\delta) = e^{-\mathrm{i}\delta\hat{H}_{odd}} e^{-\mathrm{i}\delta\hat{H}_{even}} \tag{1.48}$$

The second-order time stepper ($O(\delta^3)$ error per time step) can be obtained by symmetrizing the decomposition

$$\hat{U}^{\mathrm{TEBD2}}(\delta) = e^{-\mathrm{i}\delta/2\hat{H}_{odd}} e^{-\mathrm{i}\delta\hat{H}_{even}} e^{-\mathrm{i}\delta/2\hat{H}_{odd}}. \tag{1.49}$$

Similarly we can get the fourth-order time stepper ($O(\delta^5)$ error per time step)

$$\hat{U}^{\mathrm{TEBD4}}(\delta) = \hat{U}^{\mathrm{TEBD2}}(\delta_1)\hat{U}^{\mathrm{TEBD2}}(\delta_1)\hat{U}^{\mathrm{TEBD2}}(\delta_2)\hat{U}^{\mathrm{TEBD2}}(\delta_1)\hat{U}^{\mathrm{TEBD2}}(\delta_1) \tag{1.50}$$

with

$$\delta_1 = \frac{1}{4 - 4^{1/3}}\delta, \quad \delta_2 = \delta - 4\delta_1. \tag{1.51}$$

This strategy can be generalized to divisions of the Hamiltonian into N_h internally-commuting

summands

$$\hat{H} = \sum_{\alpha=1}^{N_h} \hat{H}_\alpha \tag{1.52}$$

and each term in $\hat{H}_\alpha$ can involve more than two sites as long as its exponential can be evaluated efficiently. We have

$$\hat{U}^{\mathrm{TEBD1}}(\delta) = e^{-\mathrm{i}\delta\hat{H}_1} e^{-\mathrm{i}\delta\hat{H}_2} \cdots e^{-\mathrm{i}\delta\hat{H}_{N_h}} \tag{1.53}$$

and

$$\hat{U}^{\mathrm{TEBD2}}(\delta) = e^{-\mathrm{i}\delta/2\hat{H}_1} e^{-\mathrm{i}\delta/2\hat{H}_2} \cdots e^{-\mathrm{i}\delta/2\hat{H}_{N_h-1}} e^{-\mathrm{i}\delta\hat{H}_{N_h}} e^{-\mathrm{i}\delta/2\hat{H}_{N_h-1}} \cdots e^{-\mathrm{i}\delta/2\hat{H}_2} e^{-\mathrm{i}\delta/2\hat{H}_1}. \tag{1.54}$$

1.7.2 MPO $W^{\mathrm{I,II}}$

Different from the TEBD, the MPO $W^{\mathrm{I,II}}$ methods do not use the Trotter-Suzuki decomposition but constructs an MPO of the approximate time evolution operator. Suppose the Hamiltonian of a length L system can be written as $\hat{H} = \sum_x \hat{H}_x$. The time evolution operator $\hat{U}(\mathrm{i}\delta)$ can be Taylor expanded as

$$e^{\delta\hat{H}} = 1 + \delta \sum_j \hat{H}_x + \delta^2/2 \sum_{xy} \hat{H}_x \hat{H}_y + \ldots. \tag{1.55}$$

The W^{I} method corresponds to applying the MPO of the approximate time evolution operator

$$\hat{U}^{\mathrm{I}}(\delta) = 1 + \delta \sum_x \hat{H}_x + \delta^2 \sum_{x \prec y} \hat{H}_x \hat{H}_y + \delta^3 \sum_{x \prec y \prec z} \hat{H}_x \hat{H}_y \hat{H}_z + \ldots, \tag{1.56}$$

where $x \prec y$ means the sites that support $\hat{H}_x$ are strictly *left* to those of $\hat{H}_y$. The first error appears in the δ^2 term. Contrary to the finding in Ref. [17], we notice that there can be far more than $O(L)$ terms omitted in δ^2. Let us consider $\hat{H} = \sum_{i=1}^{3}\sum_{j=1}^{3}\hat{S}_i^z\hat{S}_j^z$ for example. The δ^2 term in $\hat{U}^{\mathrm{I}}(\delta)$ for this Hamiltonian will be

$$(\hat{S}_1^z)^2(\hat{S}_2^z)^2 + (\hat{S}_1^z)^2(\hat{S}_3^z)^2 + (\hat{S}_2^z)^2(\hat{S}_3^z)^2 + (2\hat{S}_1^z\hat{S}_2^z)(\hat{S}_3^z)^2 + (\hat{S}_1^z)^2(2\hat{S}_2^z\hat{S}_3^z), \tag{1.57}$$

which ignores 16 contributions: 6 self-interacting terms $1/2\hat{H}_x^2$, 9 terms that $\hat{H}_x$ and $\hat{H}_y$ have overlap with each other, and 1 non-overlapping term $(2\hat{S}_1^z\hat{S}_3^z)(\hat{S}_2^z)^2$. So this approximation is generally very unstable, especially for long-range interactions. A finite state automaton can be used to construct the site tensor of $\hat{U}^{\mathrm{I}}(\delta)$

$$W_i^{\mathrm{I}} = \begin{pmatrix} \hat{1}_i + \delta\hat{D}_i & \sqrt{\delta}\hat{C}_i \\ \sqrt{\delta}\hat{B}_i & \hat{A}_i \end{pmatrix}. \tag{1.58}$$

The W^{II} method gives a better approximation to the time evolution operator. It approximates the following operator to $O(L\delta^3)$ error

$$\hat{U}^{\mathrm{II}} = 1 + \delta\sum_x \hat{H}_x + \frac{\delta^2}{2}\sum_{\langle x,y\rangle}\hat{H}_x\hat{H}_y + \frac{\delta^3}{6}\sum_{\langle x,y,z\rangle}\hat{H}_x\hat{H}_y\hat{H}_z + \ldots, \tag{1.59}$$

where $\langle x,\ldots,z\rangle$ means that for any two terms in the collection they do not cross the same bond. As an improvement of $\hat{U}^{\mathrm{I}}$, onsite terms are captured to all orders in $\hat{U}^{\mathrm{II}}$. Additional δ^2 terms show up for the example Hamiltonian[5]

$$\frac{1}{2}\left((\hat{S}_1^z)^4 + (\hat{S}_2^z)^4 + (\hat{S}_3^z)^4\right) +$$
$$(\hat{S}_1^z)^2(2\hat{S}_1^z\hat{S}_2^z) + (2\hat{S}_1^z\hat{S}_2^z)(\hat{S}_2^z)^2 + (\hat{S}_1^z)^2(2\hat{S}_1^z\hat{S}_3^z) + (2\hat{S}_1^z\hat{S}_3^z)(\hat{S}_3^z)^2 + (\hat{S}_2^z)^2(2\hat{S}_2^z\hat{S}_3^z) + (2\hat{S}_2^z\hat{S}_3^z)(\hat{S}_3^z)^2$$
$$+ (2\hat{S}_1^z\hat{S}_2^z)(2\hat{S}_2^z\hat{S}_3^z),$$

[5]There is no error to use W^{II} to construct $\hat{U}^{\mathrm{II}}$ for the example Hamiltonian, since the terms in it commutes.

(1.60)

so there are only 6 terms missing now. The site tensor takes the form

$$W_i^{\mathrm{II}} = \begin{pmatrix} \hat{W}_{D_i}^{\mathrm{II}} & \hat{W}_{C_i}^{\mathrm{II}} \\ \hat{W}_{B_i}^{\mathrm{II}} & \hat{W}_{A_i}^{\mathrm{II}} \end{pmatrix}. \tag{1.61}$$

Using auxiliary boson fields ϕ_a and $\bar{\phi}_{\bar{a}}$ ($a = 1, \ldots, n_{i-1}$ and $\bar{a} = 1, \ldots, n_i$), the blocks in W^{II} are generated by the Gaussian operator defined in the space $\mathcal{H}_\phi \otimes \mathcal{H}_{\bar{\phi}} \otimes \mathcal{H}_{physical}$

$$e^{\phi\cdot\hat{A}\cdot\bar{\phi}+\phi\cdot\sqrt{\delta}\hat{B}+\sqrt{\delta}\hat{C}\cdot\bar{\phi}+\delta\hat{D}} = \hat{W}_D^{\mathrm{II}} + \hat{W}_C^{\mathrm{II}}\cdot\bar{\phi} + \phi\cdot\hat{W}_B^{\mathrm{II}} + \phi\cdot\hat{W}_A^{\mathrm{II}}\cdot\bar{\phi} + \ldots, \tag{1.62}$$

where $\hat{W}_D^{\mathrm{II}}$ is simply the onsite term $e^{\delta\hat{D}}$. Since our approximation is to leading order, we can truncate $\mathcal{H}_{\phi_a}$ and $\mathcal{H}_{\bar{\phi}_{\bar{a}}}$ to limit maximum occupancy to be one, i.e. $\phi_a^2 = \bar{\phi}_{\bar{a}}^2 = 0$. Replacing the fields by hard-core boson creation operators

$$\phi_a \to c_a^\dagger, \quad \bar{\phi}_{\bar{a}} \to \bar{c}_{\bar{a}}^\dagger, \tag{1.63}$$

we have the operator entries in W^{II} are the transition amplitudes in $\mathcal{H}_\phi \otimes \mathcal{H}_{\bar{\phi}}$

$$\begin{aligned}
\hat{W}_{A;a\bar{a}}^{\mathrm{II}} &= \langle 0,\bar{0}|c_a\bar{c}_{\bar{a}}e^{c^\dagger\cdot\hat{A}\cdot\bar{c}^\dagger+c^\dagger\cdot\sqrt{\delta}\hat{B}+\sqrt{\delta}\hat{C}\cdot\bar{c}^\dagger+\delta\hat{D}}|0,\bar{0}\rangle \\
&= \langle 0,\bar{0}|c_a\bar{c}_{\bar{a}}e^{c_a^\dagger\bar{c}_{\bar{a}}^\dagger\hat{A}_{a\bar{a}}+c_a^\dagger\sqrt{\delta}\hat{B}_a+\sqrt{\delta}\hat{C}_{\bar{a}}\bar{c}_{\bar{a}}^\dagger+\delta\hat{D}}|0,\bar{0}\rangle \\
\hat{W}_{B;a}^{\mathrm{II}} &= \langle 0,\bar{0}|c_a e^{c_a^\dagger\bar{c}_{\bar{a}}^\dagger\hat{A}_{a\bar{a}}+c_a^\dagger\sqrt{\delta}\hat{B}_a+\sqrt{\delta}\hat{C}_{\bar{a}}\bar{c}_{\bar{a}}^\dagger+\delta\hat{D}}|0,\bar{0}\rangle \\
&= \langle 0,\bar{0}|c_a e^{c_a^\dagger\sqrt{\delta}\hat{B}_a+\delta\hat{D}}|0,\bar{0}\rangle \\
\hat{W}_{C;\bar{a}}^{\mathrm{II}} &= \langle 0,\bar{0}|\bar{c}_{\bar{a}}e^{c_a^\dagger\bar{c}_{\bar{a}}^\dagger\hat{A}_{a\bar{a}}+c_a^\dagger\sqrt{\delta}\hat{B}_a+\sqrt{\delta}\hat{C}_{\bar{a}}\bar{c}_{\bar{a}}^\dagger+\delta\hat{D}}|0,\bar{0}\rangle \\
&= \langle 0,\bar{0}|\bar{c}_{\bar{a}}e^{\sqrt{\delta}\hat{C}_{\bar{a}}\bar{c}_{\bar{a}}^\dagger+\delta\hat{D}}|0,\bar{0}\rangle.
\end{aligned} \tag{1.64}$$

In the fundamental representation of the hard-core bosons, the transition amplitudes become

$$\hat{W}_S^{\mathrm{II}} = \begin{pmatrix} \delta_{S,D} & \delta_{S,C;\bar{a}} & \delta_{S,B;a} & \delta_{S,A;a\bar{a}} \end{pmatrix} \exp \begin{pmatrix} \delta\hat{D} & 0 & 0 & 0 \\ \sqrt{\delta}\hat{C}_{\bar{a}} & \delta\hat{D} & 0 & 0 \\ \sqrt{\delta}\hat{B}_a & 0 & \delta\hat{D} & 0 \\ \hat{A}_{a\bar{a}} & \sqrt{\delta}\hat{B}_a & \sqrt{\delta}\hat{C}_{\bar{a}} & \delta\hat{D} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (1.65)$$

For a Hamiltonian that is a sum of commuting terms, we actually can utilize the auxiliary boson fields to construct the exact MPO of the time evolution operator provided that we keep the maximum occupancy to be large enough. In fact, the maximum occupancy of a boson for the ith bond is equal to one plus the number of $\hat{H}_x$ that can across that bond in a term $\prod_x \hat{H}_x$ of the Taylor expansion of the time evolution operator. Consider the example Hamiltonian again, we can actually use a larger maximum occupancy to construct a better approximation of $\hat{U}(\mathrm{i}\delta)$ than the W^{II} method. Also notice that in TEBD there is no Trotter error for a Hamiltonian that consists of commuting terms, since $e^{\sum_x \hat{H}_x} = \prod_x e^{\hat{H}_x}$. So TEBD is exact and efficient if $\hat{H}_x$ is a nearest-neighbor interaction like $\hat{S}_x^z \hat{S}_{x+1}^z$. However, for all-to-all interacting Hamiltonian like $\sum_{ij} \hat{S}_i^z \hat{S}_j^z$, we will need $O(L^3)$ number of SWAP operations which make TEBD quite inefficient although it is exact.

1.7.3 Time-dependent variational principle

The time-dependent variational principle (TDVP)[62] corresponds to project the time-dependent Schrödinger equation to the tangent space of the MPS manifold $\mathcal{M}_{|\psi[M]\rangle}$ at the current time t, i.e.

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}|\psi[M]\rangle = \hat{P}_{T\mathcal{M}_{|\psi[M]\rangle}} \hat{H}|\psi[M]\rangle. \quad (1.66)$$

The one-site tangent space projector can be decomposed as

$$\hat{P}_{T\mathcal{M}_{|\psi[M]\rangle}} = \sum_{n=1}^{N} \hat{P}^L_{n-1} \otimes \hat{1}_n \otimes \hat{P}^R_{n+1} - \sum_{n=1}^{N-1} \hat{P}^L_n \otimes \hat{P}^R_{n+1}, \tag{1.67}$$

where

$$\begin{aligned} \hat{P}^L_n &= \sum_{b_n=1}^{m_n} |\phi^L_{b_n}\rangle\langle\phi^L_{b_n}|, \\ \hat{P}^R_n &= \sum_{b_{n-1}=1}^{m_{n-1}} |\phi^R_{b_{n-1}}\rangle\langle\phi^R_{b_{n-1}}|. \end{aligned} \tag{1.68}$$

With the tangent space projector, the right hand side of Eq (1.66) becomes

$$\begin{aligned} \hat{P}_{T\mathcal{M}_{|\psi[M]\rangle}} \hat{H}|\psi[M]\rangle = &\sum_{n=1}^{N} \sum_{b'_{n-1}s'_n b'_n; b_{n-1}s_n b_n} |\phi^L_{b'_{n-1}}\rangle|s'_n\rangle|\phi^R_{b'_n}\rangle [H(n)]_{b'_{n-1}s'_n b'_n; b_{n-1}s_n b_n} [C_n]^{s_n}_{b_{n-1}b_n} \\ &- \sum_{n=1}^{N-1} \sum_{a'_n b'_n; a_n b_n} |\phi^L_{a'_n}\rangle|\phi^R_{b'_n}\rangle [K(n)]_{a'_n b'_n; a_n b_n} [D_n]_{a_n b_n}. \end{aligned} \tag{1.69}$$

Actually Eq (1.66) can be integrated into the form

$$|\psi(t+\Delta t)\rangle = \exp\left[-\mathrm{i}\hat{P}_{T\mathcal{M}_{|\psi[M]\rangle}} \hat{H}\Delta t\right]|\psi(t)\rangle. \tag{1.70}$$

The exponential operator in the right hand side can be splitted by the Lie-Trotter decomposition. To first order, it formally becomes

$$\begin{aligned} \exp\left[-\mathrm{i}\hat{P}_{T\mathcal{M}_{|\psi[M]\rangle}} \hat{H}\Delta t\right] &= \exp\left[-\mathrm{i}\sum_{n=1}^{N} \hat{P}^L_{n-1} \otimes \hat{1}_n \otimes \hat{P}^R_{n+1} \hat{H}\Delta t + \mathrm{i}\sum_{n=1}^{N-1} \hat{P}^L_n \otimes \hat{P}^R_{n+1} \hat{H}\Delta t\right] \\ &= \exp\left[-\mathrm{i}\hat{P}^L_{N-1} \otimes \hat{1}_N \hat{H}\Delta t\right] \exp\left[\mathrm{i}\hat{P}^L_{N-1} \otimes \hat{P}^R_N \hat{H}\Delta t\right] \\ &\quad \exp\left[-\mathrm{i}\hat{P}^L_{N-2} \otimes \hat{1}_{N-1} \otimes \hat{P}^R_N \hat{H}\Delta t\right] \ldots \\ &\quad \exp\left[\mathrm{i}\hat{P}^L_1 \otimes \hat{P}^R_2 \hat{H}\Delta t\right] \exp\left[-\mathrm{i}\hat{1}_1 \otimes \hat{P}^R_2 \hat{H}\Delta t\right] + \mathcal{O}(\Delta t^2). \end{aligned} \tag{1.71}$$

Higher order decomposition can be derived accordingly. The decomposition above enables us to integrate the differential equation iteratively, i.e. at step n we assume only C_n is time-dependent so we only need to solve the local effective equations

$$\begin{aligned} \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}C_n(t) &= H(n)C_n(t), \\ -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}D_n(t) &= K(n)D_n(t). \end{aligned} \tag{1.72}$$

Besides the Trotter error, there are errors from the projection to the tangent space, i.e.

$$\|(\hat{1} - \hat{P}_{T\mathcal{M}_{|\psi[M]\rangle}})\hat{H}|\psi[M]\rangle\|. \tag{1.73}$$

The projection error can be estimated according to section III. F of [63] and [64]. The key observation is that both the projection error and the Trotter error dependend on the tangent space projector. Consider the limit that the states of which the projector consists form a complete basis. Then the projector is simply an identity operator. Consequently Eq. (1.73) becomes zero and each term in the Lie-Trotter splitting commutes with each other so the Trotter error also becomes zero. In other words, $\hat{H}|\psi[M]\rangle$ is still in the same manifold $\mathcal{M}_{\mathrm{MPS}}$ so the projection does not take any effect and the equation becomes the exact Schrödinger equation and the left hand side can be exactly equal to the right hand side of (1.66). However, exactly complete basis is not reachable when the system size becomes large and we always need to compress the MPS to have a finite bond dimension. To reduce the errors, the key is to get an approximate projector of certain bond dimension as close to the identity as possible. Since $\hat{P}_{T_{|\psi[M]\rangle}\mathcal{M}_{\mathrm{MPS}}}$ originates from the variational ansatz itself, the problem becomes how to obtain a good enough basis set for the variational ansatz. Although the physical variational ansatz is always the time-evolving state which is unchangeable, the MPS representation of it can indeed be varied through the gauge freedom. For example, the MPS representation of a product state do not necessarily have bond dimension to be one but can have the maximum bond dimension allowed at each bond with the orthogonality center

being rank one, and therefore we can have complete bases for $\hat{P}^L$ and $\hat{P}^R$ even for a product state. If we re-orthogonalize the MPS, the bond dimension will shrink back to one, but if a time evolution is immediately followed at the orthogonality center, the bond dimension can be maintained.

1.8 Finite temperature

Using time evolution methods to simulate a thermal ensemble at finite temperature requires to deal with the mixed states. One can imaginary time evolve the MPO of identity[65] to get the density matrix $e^{-\beta\hat{H}}$ of a mixed state at temperature $T = 1/\beta$. Two alternative ways exist within the MPS framework based on representation of pure states, which we described below.

1.8.1 Purification

A mixed state on the Hilbert space $\mathcal{H}_P$ of a physical system P can be understood as the partial trace of the density matrix of a pure state on $\mathcal{H}_P \otimes \mathcal{H}_Q$, where the ancilla Q is a replica of P. We call

$$|\psi\rangle = \sum_{a=1}^{r} s_a |a\rangle_p |a\rangle_Q \tag{1.74}$$

a purification of

$$\rho_P = \sum_{a=1}^{r} s_a^2 |a\rangle_{PP}\langle a| = \mathrm{Tr}_Q |\psi\rangle\langle\psi|. \tag{1.75}$$

A thermal density matrix can be written as

$$\rho_\beta = Z(\beta)^{-1} e^{-\beta\hat{H}} = Z(\beta)^{-1} e^{-\beta\hat{H}/2} \hat{1} e^{-\beta\hat{H}/2} = \frac{Z(0)}{Z(\beta)} e^{-\beta\hat{H}/2} \rho_0 e^{-\beta\hat{H}/2}, \tag{1.76}$$

where ρ_0 is the thermal density matrix at infinite temperature and can be purified as

$$\rho_0 = \mathrm{Tr}_Q |\psi_0\rangle\langle\psi_0|. \tag{1.77}$$

Therefore we have

$$\rho_\beta = \frac{Z(0)}{Z(\beta)} \mathrm{Tr}_Q \left[e^{-\beta\hat{H}/2} |\psi_0\rangle\langle\psi_0| e^{-\beta\hat{H}/2} \right]. \tag{1.78}$$

A *thermofield double state* is defined as

$$|\mathrm{TFD}\rangle = \frac{1}{\sqrt{Z(\beta)}} e^{-\beta\hat{H}/2} |\psi_0\rangle. \tag{1.79}$$

The weights of a density matrix should be summed to one, so

$$1 = \mathrm{Tr}_P \rho_\beta = Z(0) \mathrm{Tr}_P \mathrm{Tr}_Q |\mathrm{TFD}\rangle\langle\mathrm{TFD}| = Z(0) \langle\mathrm{TFD}|\mathrm{TFD}\rangle \tag{1.80}$$

and thus $Z(0) = 1/\langle\mathrm{TFD}|\mathrm{TFD}\rangle$. Then the expectation value of an operator is given by

$$\langle\hat{O}\rangle_\beta = \mathrm{Tr}_P \left[\hat{O}\rho_\beta \right] = Z(0) \mathrm{Tr}_P \mathrm{Tr}_Q \left[\hat{O} |\mathrm{TFD}\rangle\langle\mathrm{TFD}| \right] = \frac{\langle\mathrm{TFD}|\hat{O}|\mathrm{TFD}\rangle}{\langle\mathrm{TFD}|\mathrm{TFD}\rangle}. \tag{1.81}$$

Now we need to determine the the purified state at infinite temperature, $|\psi_0\rangle$. All d^N weights in ρ_0 should be equal

$$\rho_0 = Z(0)^{-1} \hat{1} = Z(0)^{-1} \sum_{s_1 \ldots s_N} |s_1 \ldots s_N\rangle_{PP} \langle s_1 \ldots s_N|, \tag{1.82}$$

so $Z(0) = d^N$. The identity matrix can be factorized as tensor products of local identity matrices

$$\rho_0 = \frac{1}{d^N}\hat{1} = \bigotimes_{i=1}^{N}\left(\frac{1}{d}\hat{1}_i\right), \tag{1.83}$$

where $\hat{1}_i$ can be understood as the partial trace over the ancilla of a physical site

$$\frac{1}{d}\hat{1}_i = \sum_{s_i}\frac{1}{d}|s_i\rangle_{PP}\langle s_i| = \mathrm{Tr}_Q\left[\left(\sum_{s_i}\frac{1}{\sqrt{d}}|s_i\rangle_P|s_i\rangle_Q\right)\left(\sum_{s_i}\frac{1}{\sqrt{d}}{}_P\langle s_i|_Q\langle s_i|\right)\right], \tag{1.84}$$

i.e. the purification $|\psi_0\rangle$ should be a state that maximally entangles P and Q

$$|\psi_0\rangle = \bigotimes_{i=1}^{N}\left(\sum_{s_i}\frac{1}{\sqrt{d}}|s_i\rangle_P|s_i\rangle_Q\right) = \sum_{s_1\ldots s_N}\frac{1}{\sqrt{d^N}}|s_1\ldots s_N\rangle_P|s_1\ldots s_N\rangle_Q. \tag{1.85}$$

Notice that there is a freedom to perform unitary transformations on the ancilla which leaves the physics unchanged. This freedom can be used to reduce the entanglement[66, 67, 68, 69] of $|\mathrm{TFD}\rangle$ and in further real time evolution (e.g. to calculate Green's functions).

Naive purification can have problems at low temperature $T \to 0$. Assume the ground state energy of P is zero. When $T \to 0$, i.e. $\beta \to \infty$, the action of $e^{-\beta\hat{H}/2}$ will project $|\psi_0\rangle$ to the ground state of P. In the energy representation,

$$\begin{aligned} |\psi_0\rangle &= \frac{1}{\sqrt{d^N}}\sum_n |E_n\rangle_P|E_n\rangle_Q \\ \lim_{\beta\to\infty} e^{-\beta\hat{H}/2}|\psi_0\rangle &= \frac{1}{\sqrt{d^N}}|E_0\rangle_P|E_0\rangle_Q, \end{aligned} \tag{1.86}$$

i.e $|\mathrm{TFD}\rangle$ becomes a product state between P and Q at zero temperature. If $|E_0\rangle_P$ can be accommodated in an MPS of bond dimension D, then $|E_0\rangle_P|E_0\rangle_Q$ will need an MPS of bond dimension D^2 if we take a zigzag path in the ladder (purification of a chain), i.e. alternate P site and Q site. Then the Lanczos and SVD will have a complexity $O(D^6)$, compared to

$O(D^3)$ for the DMRG to search for a pure state $|E_0\rangle_P$. But this problem can in principle disappear, since purification is not unique and we can do a unitary transformation on Q to disentangle $|E_0\rangle_Q$ to a product state without affecting the physical observables in P.

Contrary to what have been claimed in Ref. [45], there should be no big problem for purification at high T. Although the equal Schmidt values $1/\sqrt{d^N}$ at $T=\infty$ can be evolved to be unequal but similar, this does not mean it cannot be encoded efficiently by an MPS. This can be obvious if we regard site i of P and site i of Q as one effective "supersite". Starting from a product state between the supersites at $T=\infty$, the amount of entanglement introduced by the action of $e^{-\beta\hat{H}/2}$ at short time β (high T) should be small.

Let us take as an example a spin-1/2 system of 2 P sites and 2 Q sites with $\hat{H}=3/4+\hat{S}_1\cdot\hat{S}_2$. At infinite temperature $\beta=0$, the purification

$$\begin{aligned}|\psi\rangle &= \frac{1}{2}\left(|\uparrow\uparrow\rangle_P|\uparrow\uparrow\rangle_Q+|\uparrow\downarrow\rangle_P|\uparrow\downarrow\rangle_Q+|\downarrow\uparrow\rangle_P|\downarrow\uparrow\rangle_Q+|\downarrow\downarrow\rangle_P|\downarrow\downarrow\rangle_Q\right)\\ &= \frac{1}{2}\left(|\uparrow\uparrow\rangle+|\downarrow\downarrow\rangle\right)_1\otimes\left(|\uparrow\uparrow\rangle+|\downarrow\downarrow\rangle\right)_2.\end{aligned} \tag{1.87}$$

$e^{-\beta\hat{H}/2}$ can be diagonalized as

$$e^{-\beta\hat{H}/2}=\sum_{n=0}^{3}e^{-\beta E_n/2}|n\rangle\langle n|, \tag{1.88}$$

where $E_0=0$ and $E_1=E_2=E_3=1$ with

$$\begin{aligned}|0\rangle &= \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle_P-|\downarrow\uparrow\rangle_P\right)\\ |1\rangle &= |\downarrow\downarrow\rangle_P\\ |2\rangle &= \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle_P+|\downarrow\uparrow\rangle_P\right)\\ |3\rangle &= |\uparrow\uparrow\rangle_P.\end{aligned} \tag{1.89}$$

Then

$$
\begin{aligned}
e^{-\beta\hat{H}/2}|\psi\rangle = \frac{1}{2}\Big[& e^{-\beta/2}|\uparrow\uparrow\rangle_P|\uparrow\uparrow\rangle_Q \\
& + \frac{1}{2}(1+e^{-\beta/2})|\uparrow\downarrow\rangle_P|\uparrow\downarrow\rangle_Q + \frac{1}{2}(-1+e^{-\beta/2})|\downarrow\uparrow\rangle_P|\uparrow\downarrow\rangle_Q \\
& + \frac{1}{2}(-1+e^{-\beta/2})|\uparrow\downarrow\rangle_P|\downarrow\uparrow\rangle_Q + \frac{1}{2}(1+e^{-\beta/2})|\downarrow\uparrow\rangle_P|\downarrow\uparrow\rangle_Q \\
& + e^{-\beta/2}|\downarrow\downarrow\rangle_P|\downarrow\downarrow\rangle_Q\Big],
\end{aligned}
\tag{1.90}
$$

which can be rearranged as

$$
\begin{aligned}
e^{-\beta\hat{H}/2}|\psi\rangle = \frac{1}{2}\Big[& |\uparrow\uparrow\rangle_1 \left(e^{-\beta/2}|\uparrow\uparrow\rangle_2 + \frac{1}{2}(1+e^{-\beta/2})|\downarrow\downarrow\rangle_2\right) \\
& + \frac{1}{2}(-1+e^{-\beta/2})\left(|\downarrow\uparrow\rangle_1|\uparrow\downarrow\rangle_2 + |\uparrow\downarrow\rangle_1|\downarrow\uparrow\rangle_2\right) \\
& + |\downarrow\downarrow\rangle_1 \left(\frac{1}{2}(1+e^{-\beta/2})|\uparrow\uparrow\rangle_2 + e^{-\beta/2}|\downarrow\downarrow\rangle_2\right)\Big].
\end{aligned}
\tag{1.91}
$$

To calculate the entanglement between $(PQ)_1$ and $(PQ)_2$, we only need to SVD the matrix

$$
\begin{pmatrix}
e^{-\beta/2} & 0 & 0 & \frac{1}{2}(1+e^{-\beta/2}) \\
0 & 0 & \frac{1}{2}(-1+e^{-\beta/2}) & 0 \\
0 & \frac{1}{2}(-1+e^{-\beta/2}) & 0 & 0 \\
\frac{1}{2}(1+e^{-\beta/2}) & 0 & 0 & e^{-\beta/2}
\end{pmatrix}.
\tag{1.92}
$$

At high temperature $\beta = 0.1$, the singular values are 1.927, 0.024, 0.024, 0.024. This means we can drop out the latter three Schmidt states with truncation error about 10^{-4}. This example shows that unless long tails are developed in the singular value distribution, MPS is still capable to encode the entanglement growth at small β, even though the weights in Eq. (1.90) are similar but unequal.

1.8.2 Minimally entangled typical thermal states

To avoid the problem of purification when $T \to 0$, we can use instead a method[70, 71] called *minimally entangled typical thermal states* (METTS), which is based on Monte Carlo sampling over a cleverly chosen set of states that can be represented efficiently by MPSs.

The thermal average of an operator $\hat{O}$ at temperature $1/\beta$ is given by

$$\langle \hat{O} \rangle_\beta = \frac{1}{Z(\beta)} \mathrm{Tr}\left[\hat{O} e^{-\beta \hat{H}}\right] = \frac{1}{Z(\beta)} \sum_n e^{-\beta E_n} \langle n|\hat{O}|n\rangle, \tag{1.93}$$

where $|n\rangle$ is a energy eigenstate of E_n. In Monte Carlo, it is standard to write $\langle \hat{O} \rangle_\beta$ as

$$\langle \hat{O} \rangle_\beta = \frac{1}{Z(\beta)} \sum_i \langle i|e^{-\beta \hat{H}/2} \hat{O} e^{-\beta \hat{H}/2}|i\rangle = \sum_i \frac{P_\beta(i)}{Z(\beta)} \langle \phi_\beta(i)|\hat{O}|\phi_\beta(i)\rangle = \overline{\langle \phi_\beta(i)|\hat{O}|\phi_\beta(i)\rangle}, \tag{1.94}$$

where $\{|i\rangle\}$ is a set of orthonormal basis of the system and

$$|\phi_\beta(i)\rangle = P_\beta(i)^{-1/2} e^{-\beta \hat{H}/2}|i\rangle \tag{1.95}$$

with normalization factor $P_\beta(i)^{-1/2} = \langle i|e^{-\beta \hat{H}}|i\rangle$. Since $\sum_i P_\beta(i)/Z(\beta) = Z(\beta)^{-1}\mathrm{Tr} e^{-\beta \hat{H}} = 1$, we can estimate the expectation $\langle \hat{O} \rangle_\beta$ by sampling $|\phi_\beta(i)\rangle$ with the probability $P_\beta(i)/Z(\beta)$ and calculate the average of $\langle \phi_\beta(i)|\hat{O}|\phi_\beta(i)\rangle$ in a random experiment. $|\phi_\beta(i)\rangle$ is called a METTS if $|i\rangle$ is chosen to be a classical product state

$$|i\rangle = |s_1^i\rangle |s_2^i\rangle \ldots |s_N^i\rangle, \tag{1.96}$$

which can be represented by an MPS with bond dimension $D = 1$.

To sample with the correct probability distribution at β, we can generate a Markov chain of

states with the transition probability

$$T_\beta(i \to i') = |\langle i'|\phi_\beta(i)\rangle|^2 \tag{1.97}$$

which satisfies the detailed balance

$$P_\beta(i)T_\beta(i \to i') = P_\beta(i')T_\beta(i' \to i). \tag{1.98}$$

Start from a random classical state i_1, the kth iteration of the algorithm proceeds as follows:

1. Do imaginary time evolution and normalization to $|i_k\rangle$ and get $|\phi_\beta(i_k)\rangle$.
2. Evaluate $\langle\phi_\beta(i_k)|\hat{O}|\phi_\beta(i_k)\rangle$.
3. Collapse the state $|\phi_\beta(i_k)\rangle$ to a new classical product state $|i_{k+1}\rangle$ by quantum measurements according to the transition probability $T_\beta(i_k \to i_{k+1})$.

Notice the collapse in step 3 can be done very efficiently with MPS by using the canonical form and the fact that the resulting state after a collapse is a product state of local states on all collapsed sites and the uncollapsed remainder[45]. We repeat the above steps until we get a Markov chain of enough length R. The thermal average is then approximated by

$$\langle\hat{O}\rangle \approx \frac{1}{R}\sum_{k=1}^{R}\langle\phi_\beta(i_k)|\hat{O}|\phi_\beta(i_k)\rangle. \tag{1.99}$$

Chapter 2

Density-matrix-renormalization-group study of a one-dimensional diatomic molecule beyond the Born-Oppenheimer approximation

The content of this chapter is from Ref. [42].

2.1 Introduction

The Born-Oppenheimer (BO) approximation[72] has been the starting point of solid state physics and quantum chemistry since it was first introduced in 1927. Treating the degrees of freedom of the nuclei adiabatically turns out to be a satisfactory approximation because the mass of the nucleus is more than 10^3 times of the electron mass even for the lightest atom - hydrogen.

However, the BO approximation is no longer valid for exotic systems such as the positronium molecule[73, 74, 75] which consists of two positrons and two electrons, and the emergent biexciton molecule[76] which consists of two holes and two electrons in semiconductors, because their masses are equal or nearly so. In high precision spectroscopy experiments or in systems where energy levels cross, non-adiabatic effects involving the motions of the nuclei require a theoretical treatment beyond the BO approximation[77]. Such systems are difficult to treat analytically. Various numerical approaches, such as the stochastic variational method (SVM)[78, 79, 80], quantum Monte Carlo (QMC) methods[81, 82], and Exact Factorization[77, 83, 84] combined with Density Functional Theory (DFT), have been applied to explore the spectrum of the systems in two or three dimensions and have correctly predicted the bound ground state[74] and possible bound excited states[79, 80] later proved by experiments[75].

The hydrogen molecule (H_2) and the positronium molecule (Ps_2) are in nearly opposite limits of mass ratios between the nuclei and electrons, 1836:1 vs 1:1, corresponding to adiabatic and non-adiabatic limits, respectively. Unlike H_2, for which the BO approximation can be used to simplify the numerical treatments[85], the non-adiabatic features of Ps_2 requires a complete four-body treatment. The electrons in H_2 can be in either a bonding or anti-bonding state, corresponding to a spin singlet or triplet respectively, and the anti-bonding state is unstable against dissociation into two atoms. There are also two types of nuclear spin states, called *spin isomers*, with the singlet known as para-hydrogen and the triplet known as ortho-hydrogen. In Ps_2, if both the electrons and positrons are in spin singlet states, the molecule is bound, while the triplet-triplet excited state is unbound[78, 79, 80, 86, 87]. Similar behavior is found for the biexciton, which has a typical mass ratio $m_e/m_h = 0.67$. Therefore, a crossover where the spin state of the "nuclei" starts to influence the binding of the molecule should exist when one tunes the mass ratio from that of H_2 to that of Ps_2, corresponding to the breakdown of the BO approximation.

Recently, Fisher and Radzihovsky have argued that nuclear spin can cause significant changes in chemical reactions even at room temperature[88]. In this article, we use the density matrix renormalization group (DMRG) method[1, 89] to study a 1D version of H_2 with mass ratio $1 \leq m_p/m_e \leq 1000$ with high precision[1]. While systems with four quantum particles have previously been studied for 2D and 3D, our technique can easily extend to dozens of 1D particles, beyond the reach of many 2D and 3D techniques.

Using DMRG, we are able to find the ground state of a one dimensional fermionic four-body system, i.e. the diatomic molecule with tunable mass ratio, and measure its physical observables such as the ground state energy, density-density correlation, and entanglement between particles. In the regime of mass ratio $m_p/m_e \gg 1$ as a benchmark, the results match the BO approximation, as expected. At mass ratio $m_p/m_e = 1$, our results match the behavior of 3D Ps_2: its singlet-singlet four-body ground state is bound while the triplet-triplet state is unbound in 1D. However, contradicting with previous SVM results in 3D[87, 86, 80], the triplet-singlet state is unbound in 1D. (Note that it is not an eigenstate of Ps_2 because of the requirement of symmetry of charge conjugation. Our nuclei and electrons are always distinguishable particles). We find that the mass ratio where these unbound states become bound is $m_p/m_e = 2.73$ for our chosen parameters of the interaction, while the singlet-singlet state is bound for all the mass ratios. Obtaining the energies and the average separations of nuclei at different mass ratios m_p/m_e of the singlet-singlet state and the triplet-singlet state, we study the passage between the adiabatic and non-adiabatic limit.

The outline of this article is as follows: first, we will introduce the microscopic model and explain the numerical techniques; then, the results from our DMRG calculations will be illustrated and a comparison to the Hartree-Fock mean field calculation will be made; finally, we will discuss the potential of our method to be used in other 1D few-body systems and many-body systems.

[1]We did not remove the center of mass motion because it will lead to additional coupling terms in the Hamiltonian[84, 90, 87].

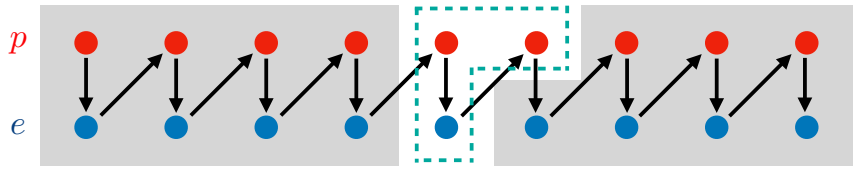

Figure 2.1: Ladder arrangement of grid points describing the discretized 1D systems. A DMRG sweep is along the zig-zag route (black arrows), where the red grid points in the upper leg represent the lattice sites for the nuclei and the blue in the bottom leg represents the electrons. The total number of sites is $2N_L$, where $N_L = L/\Delta x$, L is the size of the 1D system, and Δx is the grid spacing. The dashed outline shows three adjacent sites grouped together as part of the three-site DMRG algorithm.

2.2 Model

The Hamiltonian for a 1D system of interacting spin-$\frac{1}{2}$ nuclei ("protons" with coordinates X_i and mass m_p) and electrons (with coordinates x_i and mass m_e) is given by

$$H = -\sum_{i=1}^{N_e} \frac{1}{2m_e} \frac{\mathrm{d}^2}{\mathrm{d}x_i^2} - \sum_{i=1}^{N_p} \frac{1}{2m_p} \frac{\mathrm{d}^2}{\mathrm{d}X_i^2} + \sum_{i \geq j} V(x_i - x_j) + \sum_{i \geq j} V(X_i - X_j) - \sum_{ij} V(x_i - X_j), \quad (2.1)$$

where the spin index has been omitted. N_e and N_p are the total number of electrons and nuclei respectively. For our H_2-like diatomic system, we have $N_e = 2$ and $N_p = 2$. V is the "Coulomb" interaction whose form will be given in the next section, with the intra-species interactions being repulsive and inter-species interactions being attractive. We use atomic units, so $\hbar = 1$ and $e = 1$. The mass of the particle is measured in units of m_e, so if we denote the mass ratio $m_p/m_e = M$, then $m_e = 1, m_p = M$.

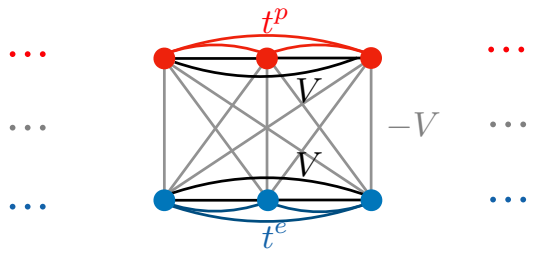

Figure 2.2: Illustration of the hopping and the interaction parameters. t^p and t^e are the hopping parameters which can be the nearest-neighbor or the next-nearest-neighbor for the nuclei and electrons respectively.

2.3 Numerical Techniques

We need first to discretize the continuous system into a lattice in order to use DMRG to study it. First, we write the Hamiltonian (2.1) in second quantized form in terms of field operators

$$H = \int \mathrm{d}x \phi^\dagger_{\alpha,s}(x) \left[-\frac{1}{2m_\alpha} \frac{\mathrm{d}^2}{\mathrm{d}x^2} \right] \phi_{\alpha,s}(x) + \frac{1}{2} \iint \mathrm{d}x \mathrm{d}x' V_{\alpha\beta}(x-x') \phi^\dagger_{\alpha,s}(x) \phi^\dagger_{\beta,s'}(x') \phi_{\beta,s'}(x') \phi_{\alpha,s}(x), \quad (2.2)$$

where $\alpha, \beta \in \{p, e\}$, $s, s' \in \{\uparrow, \downarrow\}$, and $V_{\alpha\beta} = V$ if $\alpha = \beta$ and $V_{\alpha\beta} = -V$ if $\alpha \neq \beta$. The Einstein summation convention has been used. The field operators satisfy the canonical anti-commutation relation for fermions

$$\{\phi^\dagger_{\alpha,s}(x), \phi_{\beta,s'}(y)\} = \delta(x-y)\delta_{\alpha\beta}\delta_{ss'}.$$

Notice that we choose the interspecies operators to anti-commute. This does not matter as long as we keep the different species of particles distinguishable in the implementation.

The discretization of the field operator $\phi^\dagger_{\alpha,s}(x = x_i = (i - \frac{1}{2})\Delta x) = \frac{1}{\sqrt{\Delta x}} c^\dagger_{i,\alpha,s}$ with grid

spacing Δx can be understood as an approximate canonical transformation

$$\phi^{\dagger}_{\alpha,s}(x) \approx \sum_{i=1}^{N_L} \psi_i(x) c^{\dagger}_{i,\alpha s} \tag{2.3}$$

with the orthonormal (approximately complete) basis set being

$$\psi_i(x) = \begin{cases} \dfrac{1}{\sqrt{\Delta x}}, & \text{if } \ (i-1)\Delta x \le x < i\Delta x; \\ \\ 0, & \text{otherwise,} \end{cases} \tag{2.4}$$

and thus the the density operator being

$$\rho_{\alpha,s}(x) \equiv \phi^{\dagger}_{\alpha,s}\phi_{\alpha,s} \approx \frac{1}{\Delta x} c^{\dagger}_{i,\alpha s} c_{i,\alpha s} \equiv \frac{1}{\Delta x} n_{i,\alpha s} \tag{2.5}$$

for $(i-1)\Delta x \le x < i\Delta x$.

Combined with the fourth order finite-difference formula for the second derivative

$$\frac{\mathrm{d}^2\phi(x)}{\mathrm{d}x^2} = \frac{1}{12(\Delta x)^2}\big[-\phi(x+2\Delta x) + 16\phi(x+\Delta x) \\ - 30\phi(x) + 16\phi(x-\Delta x) - \phi(x-2\Delta x)\big] + \mathcal{O}((\Delta x)^4),$$

the Hamiltonian is discretized to be

$$H = \sum_{i,\alpha} t_0^{\alpha} n_{i,\alpha} + \sum_{\langle i,j\rangle,\alpha s} t_1^{\alpha} c^{\dagger}_{i,\alpha s} c_{j,\alpha s} + \sum_{\langle\langle i,j\rangle\rangle,\alpha s} t_2^{\alpha} c^{\dagger}_{i,\alpha s} c_{j,\alpha s} \\ + \sum_{i,\alpha} V(0) n_{i,\alpha\uparrow} n_{i,\alpha\downarrow} - \sum_{i} V(0) n_{i,p} n_{i,e} + \sum_{i>j,\alpha\beta} V_{ij}^{\alpha\beta} n_{i,\alpha} n_{j,\beta}, \tag{2.6}$$

where $n_{i,\alpha} = n_{i\alpha\uparrow} + n_{i\alpha\downarrow} = \Delta x \sum_s \rho_{\alpha,s}(x_i) \equiv \Delta x \sum_s \phi^{\dagger}_{\alpha,s}(x_i)\phi_{\alpha,s}(x_i) = \sum_s c^{\dagger}_{i,\alpha s} c_{i,\alpha s}$, $t_0^{\alpha} = \frac{5}{4\eta}$, $t_1^{\alpha} = -\frac{2}{3\eta}$, $t_2^{\alpha} = \frac{1}{24\eta}$, with $\eta \equiv m_{\alpha}(\Delta x)^2$, and $V_{ij}^{\alpha\beta} = V_{\alpha\beta}((i-j)\Delta x)$. Notice that now $1 \le i, j \le N_L$ label the site points. To fourth order in Δx, only hoppings up to next-nearest

neighbor remain. The discretization actually introduces a length scale Δx into the system, which corresponds to a cutoff in the energy scale. The larger the Δx is, the lower is the energy cutoff, which means we truncate the Hamiltonian to lower energy sectors. There is an upper bound for Δx, since it can not exceed the wavelength of the low-energy excitations. $N_L = L/\Delta x$ actually limits the maximum number of particles we can put into the system, which has similar function to the maximum number of bands in the Hubbard model within the BO approximation. For the molecule, we use a grid spacing $\Delta x = 0.1$, which we find is accurate for energies to a relative error of about 10^{-4}.

To accommodate the two oppositely charged species of particles, the geometry of the system is represented by a two-leg ladder(FIG. 2.1), with each species living in one of the legs. Hopping is only along the legs and the interactions can be either along the legs (repulsive) or between the legs (attractive).

Now we explain the form of the Coulomb interaction V we use. The $1/x$ form of the Coulomb potential in 1D is numerically difficult and unphysical because of its singularity at $x = 0$. Instead, there are some conventional choice for one dimensional systems, e.g. the soft Coulomb potential $1/\sqrt{x^2 + a^2}$, which is still long ranged and has no singularity at the origin if $a \neq 0$. If we are only concerned about short-range properties, an exponential form can well approximate the long-range potential and meanwhile reduce the computational complexity[91]. Therefore as a convenient choice, here we use a exponential potential of the form[91]

$$V(x) = A \exp(-\kappa|x|), \tag{2.7}$$

where $A = 1.071295$ and $\kappa^{-1} = 2.385345$ have been shown to optimally approximate the soft Coulomb potential with $a = 1$ at short range[91]. This exponential potential nicely mimics some three dimensional electronic properties[91]. In our work, A and κ are also varied to see

their influence on the results.

To use DMRG in the two dimensional ladder system, we take as usual the zig-zag path to form a one dimensional Matrix Product State (MPS), i.e. the $p-$leg being the odd sites and the $e-$leg being the even sites. In such a way, there is no hopping between nearest neighbors, i.e. a $p-$site and a $e-$site, so the number of particles in each block cannot readily fluctuate in a conventional 2-site DMRG sweep and the optimization will get stuck. We could introduce a special noise term in the Hamiltonian to solve this problem[59]. Here, instead, we use a 3-site algorithm which naturally fits the hopping structure of the system and introduces "communication" between the next-nearest neighbors at each 3-site local update. At each local update, a singular value decomposition (SVD) is done once only at the left bond of the 3 sites for a left-to-right half-sweep, or the right bond for a right-to-left half-sweep. The computational complexity comes mainly from applying the Matrix Product Operator (MPO) to the MPS in the mixed canonical form[45]. For the two-site algorithm, the complexity is $\mathcal{O}(D^3 D_W d^2 + D^2 D_W^2 d^3)$, where D, D_W, d are respectively the bond dimension of the MPS, MPO, and the dimension of the local Hilbert space at each site; for the three-site algorithm, the complexity is $\mathcal{O}(D^3 D_W d^3 + D^2 D_W^2 d^4)$. So the complexity of the three-site algorithm is about $\mathcal{O}(d)$ times of that of the two-site one, which is acceptable. For the singlet-singlet state in a grid of $L = 40$, the number of states m needed to to achieve a truncation error of 10^{-10} is about 70, and the number of sweeps needed to reach energy convergence with error smaller than 10^{-6} is about 160 (see FIG. 2.3). The large number of sweeps needed is due to the fine grid spacing and large associated kinetic energy scale $1/(\Delta x)^2$.

To accelerate the calculation, we utilize a compression algorithm[35] which uses singular value decompositions (SVDs) to reduce the bond dimension of the MPO. The factorizability of the exponential function

$$V_{ij} = \lambda^{-|i-j|} = \lambda^{-i}\lambda^{j} \quad (i > j) \tag{2.8}$$

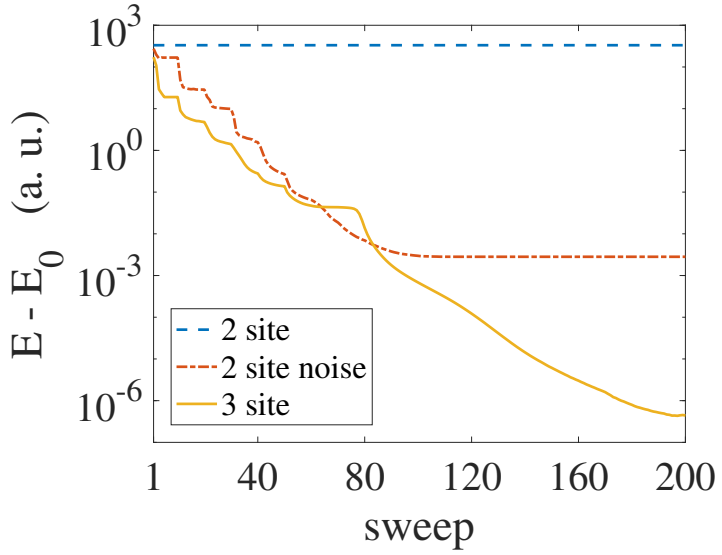

Figure 2.3: Comparison of the performance of 2-site algorithm with noise, without noise, and the 3-site algorithm (without noise). Data are taken from a DMRG simulation of system in the singlet-singlet state with $M = 3$, $L = 40$, $\Delta x = 0.1$ with 200 sweeps. Here E_0 is the converged ground state energy calculated by DMRG after 240 sweeps.

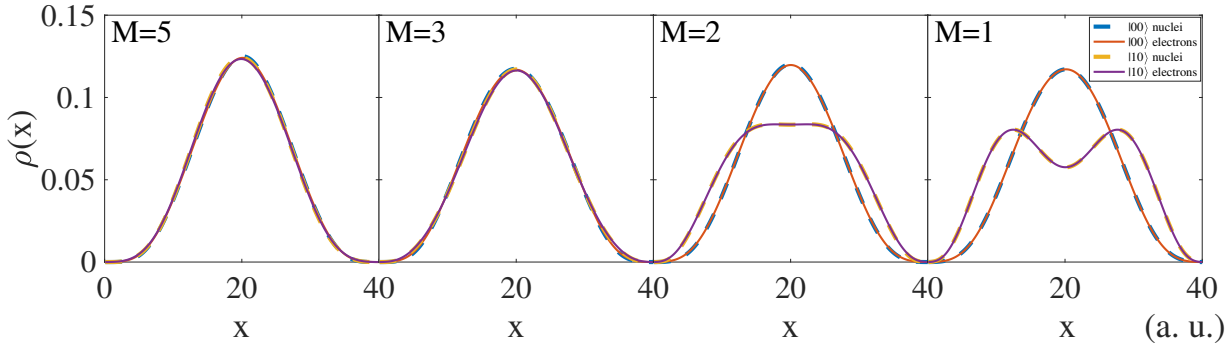

Figure 2.4: Comparison of density of particles in the singlet-singlet and the triplet-singlet ground states at different M, where the bold dash lines are for the nuclei and the thin solid lines are for the electrons. The states are labeled as $|S_p, S_e\rangle$, where $|S_p\rangle$ is the total spin of the nuclei and $|S_e\rangle$ is the total spin of the electrons. The box size is 40.

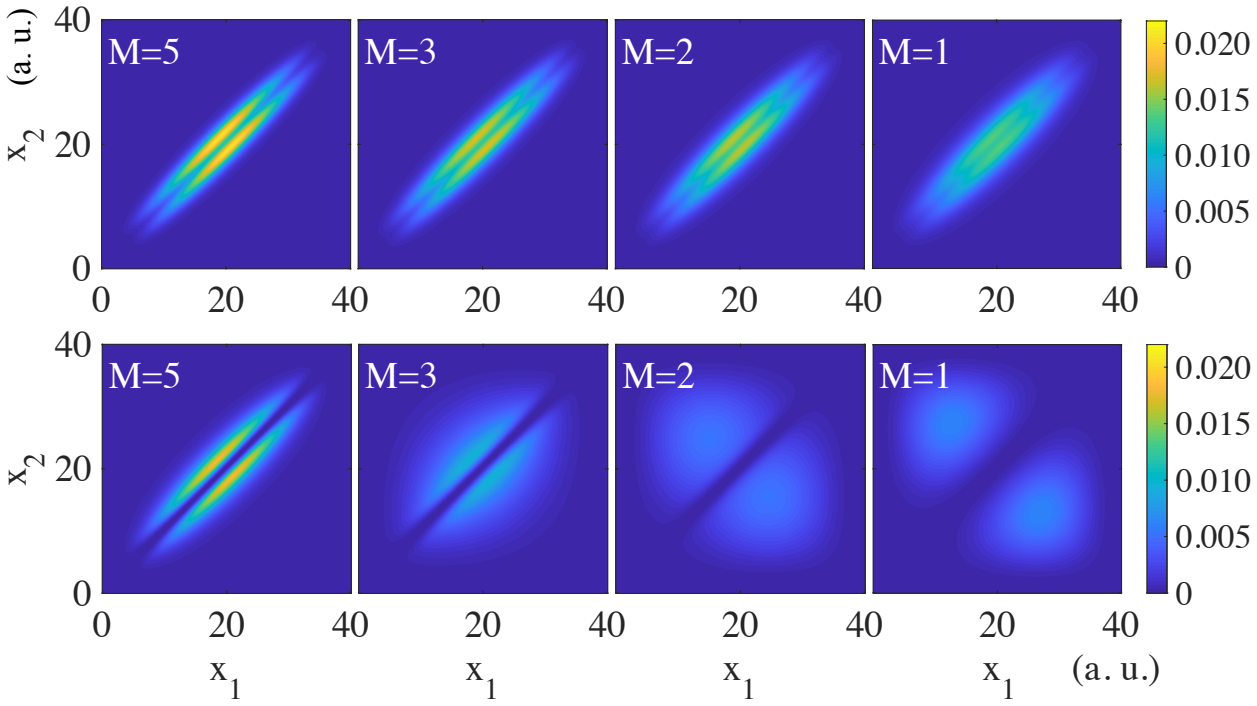

Figure 2.5: Comparison of density-density correlations of nuclei $\langle \rho_p(x_1)\rho_p(x_2)\rangle$ in the singlet-singlet (upper) and the triplet-singlet (bottom) ground states at different M.

indicates its MPO can be maximally compressed by SVDs. Other forms of long-range interactions can be expressed in terms of a sum of exponentials and the number of significant singular values is still controllable[35].

Unlike the Ps_2 molecule, which has a charge conjugation symmetry between the electron and positron, the nuclei and electrons in our system are distinguishable particles and the total spin S of each species should be conserved individually. Instead of dealing with the implementation of the global $SU(2)$ symmetry[92], a S^2_{tot} operator for species of particles in the singlet state is added to the Hamiltonian for optimization in order to achieve its conservation.

Errors of our calculation can come from: 1) discretization of the continuous system with a grid spacing $\Delta x = 0.1$; 2) finite size effects of order $\pi^2/4(M+1)L^2$ for the energy; 3) DMRG truncation errors of order 10^{-10}; 4) errors from incomplete convergence in the number of sweeps, which are about 10^{-5}.

2.4 Results

By measuring the density of particles (FIG. 2.4) and the density-density correlations of the nuclei (FIG. 2.5), keeping the electrons in the singlet state, we find that the triplet nuclei system gradually becomes unbound when we decrease the mass ratio M from 5 to 1 while the singlet nuclei system is always bound.

To characterize the binding of the molecule quantitively, we define the average separation of the nuclei d, i.e.

$$d = \sqrt{\frac{\sum_i (x_i - x_c)^2 \rho(x_i, x_c)}{\sum_i \rho(x_i, x_c)}}, \tag{2.9}$$

where $x_i = i\Delta x$, x_c is the center site, and $\rho(x_i, x_c) = \langle \Phi | \rho_p(x_i) \rho_p(x_c) | \Phi \rangle$ is the density-density correlation for the nuclei in ground state $|\Phi\rangle$, and the binding energy E_{bind}, i.e.

$$E_{\text{bind}} = E(2) - 2E(1), \tag{2.10}$$

where $E(1)$ is the ground state energy of one atom consisting of one electron and one nucleus and $E(2)$ is the ground state energy of the diatomic molecule.

From now on, we denote the triplet-singlet state as $|10\rangle$ and the singlet-singlet state as $|00\rangle$. For the $|10\rangle$ state, the average separation d of nuclei scales linearly with the box size L approaching $M = 1$, which indicates that $d \to \infty$ as $L \to \infty$ at small M, i.e. the system is unbound at small mass ratios. The error of the binding energy E_{bind} of the diatomic molecule due to finite-size effects can be estimated by the ground state energy of a particle in a box, $\pi^2/4(M+1)L^2$. If we use a system size of $L = 120$, the error is of order 10^{-4} even for the smallest mass ratio $M = 1$, which is negligible. From the data of systems of length $L \geq 120$, it is roughly observed that the binding energy is positive when $M = 3$ but approaching 0^- when $M \leq 2.5$, which means that there should be some critical mass

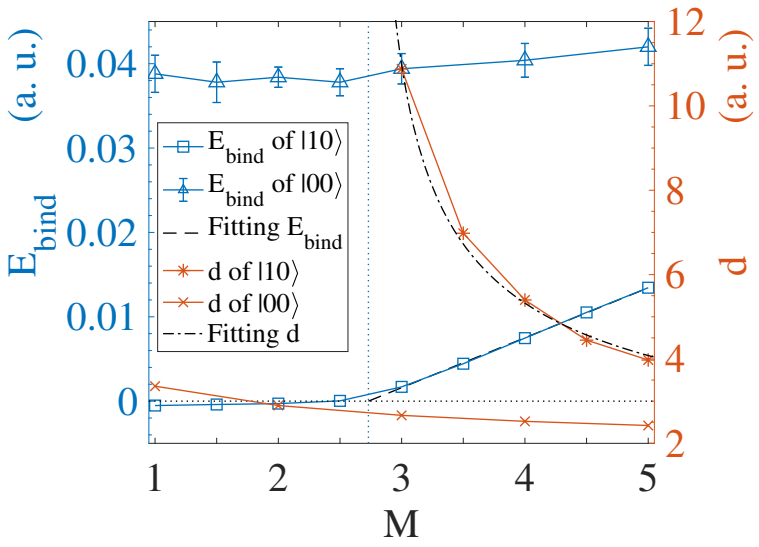

Figure 2.6: Binding energy E_{bind} and average separation d versus mass ratio M. The critical mass ratio where the molecule just binds is $M_c = 2.731$ by extrapolation. Both the data of the binding energy E_{bind} (blue square) and the average separation d (red star) for the triplet-singlet state $|10\rangle$ are taken from systems of $L = 120$, while for the singlet-singlet state $|00\rangle$ the data of d are from systems of $L = 40$ and the data of E_{bind} are from extrapolation to infinite size L. The fitting curves are $E_{\text{bind}} = a(M - M_c)$ and $d = b/\sqrt{M(M - M_c)} + d_{BO}$, where $a = 0.005918$, $b = 8.456$, and $d_{BO} = 1.571$.

ratio between 2.5 and 3 where the system changes from bound to unbound. To give an upper bound on the value of the critical mass ratio, we extrapolate E_{bind} from the bound side to get the critical mass ratio $M_c = 2.731$, which is consistent with the divergence of d approaching M_c from the right side, as shown in FIG. 2.6. This divergent behavior of d near M_c can also be fitted. Near unbinding, the size of the bound state becomes much larger than the exponential potential's decay length, so the potential becomes irrelevant and the scaling of the binding energy is only related to the kinetic energy, i.e. $E_{\text{bind}} \sim 1/Md^2$ or $d \sim 1/\sqrt{E_{\text{bind}}M}$. Combined with the extrapolation formula $E_{\text{bind}} = a(M - M_c)$, where $a = 0.005918$, we get the fitting formula for d near M_c is $d = b/\sqrt{M(M - M_c)} + d_{BO}$, where $b = 8.456$ and $d_{BO} = 1.571$. It accurately[93] predicts d_{BO}, which is the separation of the nuclei in the BO limit $M \to \infty$. For the $|00\rangle$ state, by observing its binding energy E_{bind} and the average separation d of the nuclei, we can conclude that it always binds.

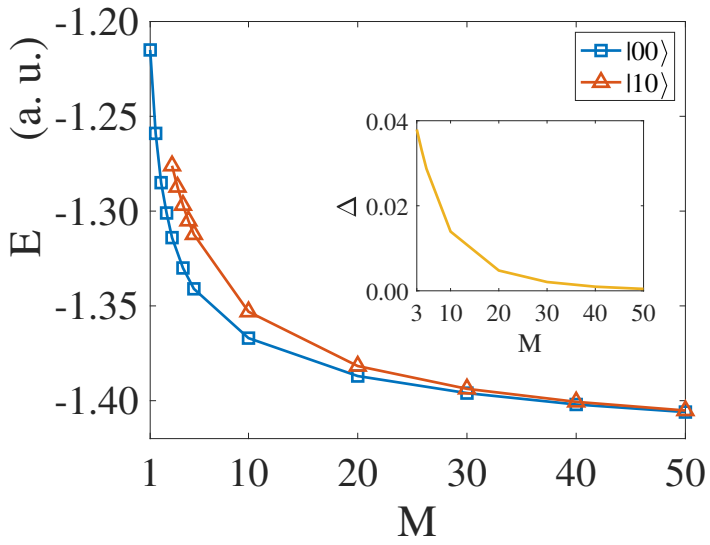

Figure 2.7: Comparison of the energy of the triplet-singlet state $|10\rangle$ and the singlet-singlet state $|00\rangle$ at different mass ratio M. Inset: the energy difference Δ between the triplet-singlet state $|10\rangle$ and the singlet-singlet state $|00\rangle$ at different mass ratio M. Data for $|00\rangle$ are from extrapolation to $L = \infty$ with error of order 10^{-3} and for $|10\rangle$ from simulation of system of size $L = 120$. Data for $|10\rangle$ when $M < 3$ has been excluded since the molecule becomes unbound.

As we mentioned before, many studies have shown that in 3D the $|00\rangle$ ground state of Ps_2 is bound and the $|11\rangle$ is unbound, with which our results at $M = 1$ in 1D are consistent. However, while they predicted the $|10\rangle$ excited state is bound in 3D, we conclude in 1D it is unbound.

In FIG. 2.7, we show the energy of the $|00\rangle$ and $|10\rangle$ states at different mass ratio M. The energy gap Δ between the two closes to 10^{-4} when M is increased to 50, where the influence of nuclei's spin on the binding energy is negligible.

The binding of the molecule can also be qualitatively illustrated in the adiabatic potential energy surface (PES) $E^e(\boldsymbol{X})$. Under the BO approximation, it is obtained by solving the clamped-nuclei Schrödinger equation

$$H^e(\boldsymbol{X})\chi_{n,\boldsymbol{X}}(\boldsymbol{x}) = E^e_n(\boldsymbol{X})\chi_{n,\boldsymbol{X}}(\boldsymbol{x}) \tag{2.11}$$

for each fixed configuration of nuclei $\boldsymbol{X} = (X_1, ..., X_{N_p})$, where $\boldsymbol{x} = (x_1, ..., x_{N_e})$ is the coordinate of the electrons and $H^e(\boldsymbol{X})$ is the Hamiltonian after separating the nuclei's kinetic part of the full Hamiltonian H, i.e.

$$\begin{aligned} H &= T^p + H^e(\boldsymbol{X}) \\ H^e(\boldsymbol{X}) &= V^{pp}(\boldsymbol{X}) + T^e + V^{ee} + V^{pe}(\boldsymbol{X}) \end{aligned} \tag{2.12}$$

with the nuclei fixed to certain configuration $\boldsymbol{X}$. This separation can only be done when $M \gg 1$ and no level crossing happens for the PES of different energy levels E_n^e so that the nuclei are almost stationary compared to electrons and the adiabatic theorem is valid. Nevertheless, for diatomic molecule at small mass ratio, we can still give an effective definition of the PES:

$$\mathcal{E}^e(R) = \frac{\langle \Phi' | H^e | \Phi' \rangle}{\langle \Phi' | \Phi' \rangle}, \tag{2.13}$$

where

$$|\Phi'\rangle = \hat{\rho}_p(x_c + R/2)\hat{\rho}_p(x_c - R/2)|\Phi\rangle \tag{2.14}$$

is the state after successively measuring (projecting) the density of nuclei at $x_c + R/2$ and $x_c + R/2$ in the eigenstate $|\Phi\rangle$ of H (here $|\Phi\rangle$ is the ground state calculated by DMRG). This measurement projects $|\Phi\rangle$ to the Hilbert subspace that has one nucleus at $x_c - R/2$ and the other one at $x_c + R/2$. When $M \gg 1$, $\mathcal{E}^e$ is equivalent to E^e in the BO approximation, as illustrated in FIG. 2.8. At smaller M, however, $T^p + \mathcal{E}^e$ is only part of an effective nuclear Hamiltonian and feedback from the nuclei's motion needs to be taken into consideration[84, 94]. Nevertheless, we can still infer some information from FIG. 2.8 about the binding of the molecule at small M.

For the $|10\rangle$ state, the overlap between the curves obtained from the BO approximation and

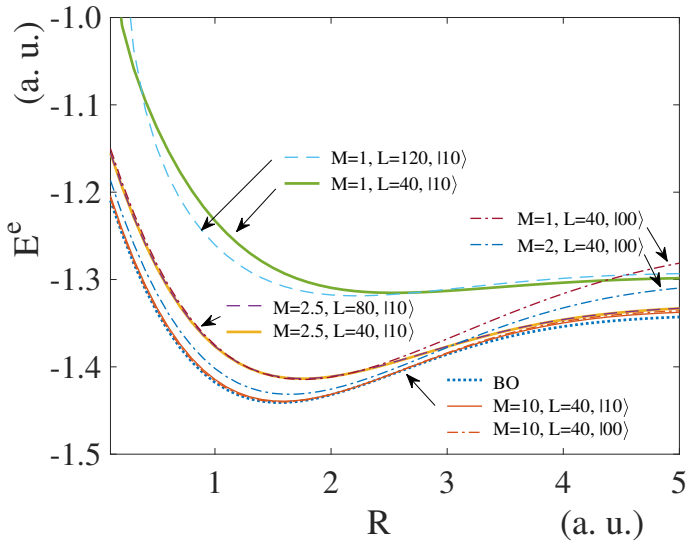

Figure 2.8: Potential energy surfaces (PES) at different mass ratio and box size. $M = m_p/m_e$ is the mass ratio and L is the box size. $|10\rangle$ and $|00\rangle$ denotes the triplet-singlet state and singlet-singlet state respectively. BO denotes the PES in the Born-Oppenheimer approximation.

from DMRG when $M \geq 10$ implies that the BO approximation works pretty well in that regime. For $M \sim 1$, the depth of the PES decreases and the minimum of the PES moves farther away from the equilibrium position of the BO approximation. Considerable finite size effect appears when $M = 1$, which can be seen by comparing the curves before and after increasing the box size. These two qualitative facts indicate that the molecule in $|10\rangle$ might be unbound when $M \sim 1$, although it should not be conclusive since $\mathcal{E}^e$ defined by Eq. (2.13) ignores part of the non-adiabatic effects from the motion of the nuclei.

For the $|00\rangle$ state, the curves coincide with that of the $|10\rangle$ state when $M > 10$, which indicates in that regime the spin of the nuclei does not affect the binding of the molecule and can be treated classically. When $M = 1$, however, the PES of the $|00\rangle$ state differs from that of the $|10\rangle$ state by being much deeper and having a minimum closer to the origin, which verifies the binding nature of the $|00\rangle$ state.

By tuning the parameters A and κ of the exponential potential and using other forms such as the soft-Coulomb or rounded exponential $V(x) = A\exp(-\kappa\sqrt{x^2 + 1/4})$ (not illustrated

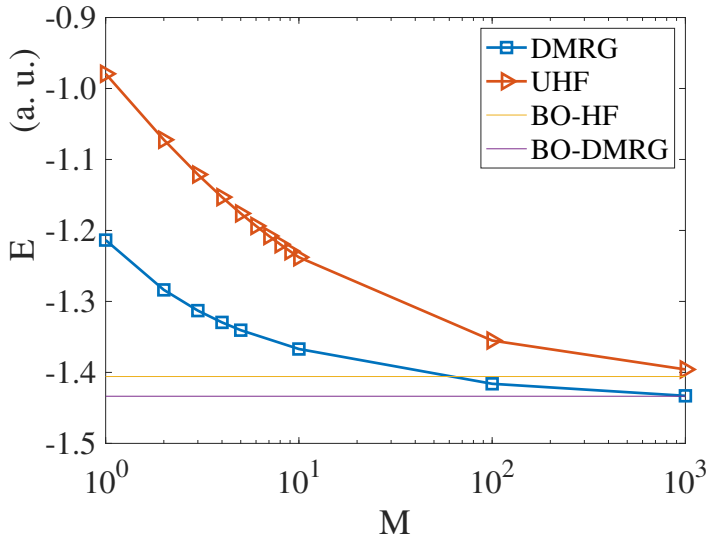

Figure 2.9: Energy of the singlet-singlet state $|00\rangle$ versus mass ratio for unrestricted Hartree-Fock (UHF) calculation beyond the BO approximation. Also shown are the restricted and unrestricted Hartree-Fock calculation under BO approximation (BO-HF), both of which give the same energy, and DMRG energy under BO approximation (BO-DMRG). The separation between nuclei we used under BO is $R = 1.6$. The kinetic energy of the nuclei's relative oscillation is added to the BO energy so as to compare with the energy beyond BO. Data are all taken from system of $L = 40$.

here), we find that $|00\rangle$ is always bound for all M independent of the specific form of the interaction. For $|10\rangle$, the critical mass ratio M_c where the molecule becomes unbound is changed with the shape of the potential, i.e. κ and A, and the form of the potential.

We also investigated the case of spinless bosonic nuclei, which turns out to be equivalent to the singlet fermion nuclei case because they have the same symmetry requirement for the spatial part of the wavefunction.

In a molecule, the Hartree-Fock (HF) approximation is often a good starting point. However, without the BO approximation, the separation between an electron and a nucleus appears as a two-particle correlation, rather than a single-particle effect. This fact makes a simple generalization of HF a poor approximation, which is illustrated in FIG. 2.9, where we did unrestricted Hartree-Fock mean field calculations both within and beyond the BO approximation in a discretized grid for the diatomic molecule as a comparison. Unlike HF under

BO, which includes the nuclei's interaction with electrons by introducing an external potential after fixing the position of the nuclei at the equilibrium positions and optimizes the electrons' orbitals, our non-BO UHF ansatz of the whole diatomic molecule is a factorization into Slater determinants of electrons and nuclei, where the single-particle wavefunctions of both species are optimized.

At the large mass ratio $M = 10^3$, the energy of the BO-DMRG and DMRG calculations agree quite well. Correlations result in an expected small energy difference between BO-DMRG and BO-HF. Perhaps less expected is a small but noticeable disagreement between the BO-HF and non-BO HF calculations. While the BO-HF gives a satisfactory approximation of electrons' wavefunction in the BO limit, the non-BO UHF assumption to factorize the wavefunction of the whole molecule into the electrons and nuclei's parts fails because of the attractive nature of the interaction and the non-adiabatic movement of the electrons with the nuclei at small M, as illustrated by the large discrepancy between the non-BO UHF and DMRG at small M in FIG. 2.9. To explain this point, let us consider the simpler case of a single hydrogen atom with the mass ratio M being tuned, where we do not change to center of mass or relative coordinates (since this is much less useful for our discussion of the molecule). In this case BO-HF is exact at $M \to \infty$, since the wavefunction is single-particle, i.e. $\phi(x)$, where x is the electron's coordinate; non-BO HF at small M is not exact, since it approximates the wavefunction of the whole atom $\phi(x, X)$ as the product of two orbitals, $\psi(x)\chi(X)$, where x and X are the electron's and nucleus's coordinates respectively. As we mentioned before, it indicates that while the single-particle picture works well for electrons at large M when the BO approximation is valid, it fails to predict the correct behavior of the four-body system at small M.

Chapter 3

Time-dependent variational principle with ancillary Krylov space

The content of this chapter is from Ref. [44].

3.1 Introduction

At the heart of the success of the density matrix renormalization group (DMRG)[1, 89] for approximating the ground states of one- and two-dimensional lattice systems is the matrix product state (MPS) representation underlying it. MPS are ideal for one-dimensional gapped ground states[5, 6], and are also a powerful approximation for one-dimensional gapless ground states and ground states of finite-width two-dimensional[95] cylinders and strips.

MPS are also useful in solving the *time-dependent* Schrödinger equation. Vidal's time-evolving block decimation (TEBD) method[15, 16, 19] can be framed as a slight change to the DMRG sweeping algorithm, called the time-dependent DMRG method (tDMRG)[61]. Since the invention of TEBD, a number of variations have been developed, and have been

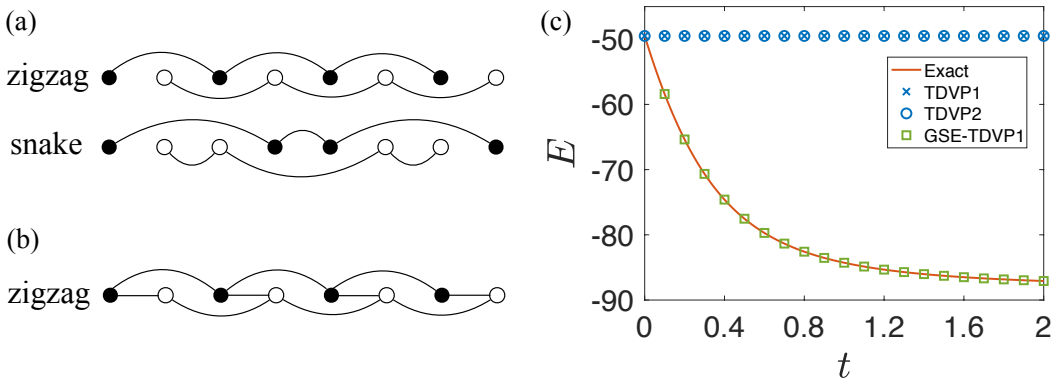

Figure 3.1: (a) Interactions in the rung-decoupled Heisenberg ladder after mapping the lattice to a one-dimensional geometry by using zigzag or snake path. Black(white) circles are sites in the first(second) leg. Only circles connected by lines have interactions between them. (b) Interactions in the rung-coupled Heisenberg ladder. (c) Failure of the one-site TDVP (TDVP1) and two-site TDVP (TDVP2) for the imaginary time evolution of the rung-decoupled Heisenberg ladder of leg length 100. GSE-TDVP1 is our method. The zigzag path is used.

widely used, for example in computing spectral functions[96] and in simulating the dynamics of cold atom systems[97].

Treating the time evolution of systems with long-range interactions is more difficult. In its original form, TEBD handles only nearest-neighbor interactions. A simple modification, exploiting swap gates to move sites which are not next to each other to be temporarily adjacent, is effective for systems with only a modest number of beyond-nearest-neighbor interactions. Alternatively, an approach[17] to approximate the time evolution operator in terms of a matrix product operator (MPO) was developed to treat long-range interactions efficiently, and has proven successful in calculating the response function of the spin-1/2 Haldane-Shastry model[17]. Another approach[43] allowing the simulation of dynamics with long-range interactions is based on Runge-Kutta, and an improved version of it was recently found useful in treating the dynamics of chemical systems[98].

Some of the most attractive MPS time evolution methods which can deal with long-range interactions are based on the time dependent variational principle (TDVP)[99, 18], where the time-dependent Schrödinger equation is projected to the tangent space of the MPS

manifold of fixed bond dimension at the current time. The manifold of all possible MPS's with a particular bond dimension can be thought of as a constraining surface for the time evolution, which in the case of real time can be thought of analogously to classical motion under a constraint. As in classical mechanics, which possesses a symplectic structure, the probability and energy is automatically conserved, and so are the other integrals of motion provided that the corresponding symmetry transformation does not take the state out of the manifold[63]. Energy and probability are not conserved in most other time dependent MPS methods due to the necessity of truncation to keep an efficient MPS representation.

The initial implementation[18] of TDVP simultaneously updates all MPS tensors and suffers from numerical instability. The difficulties were overcome later by an alternative integration scheme[62] based on a Lie-Trotter decomposition of the tangent space projector. This algorithm, in the imaginary time case, as the time step goes to infinity, is equivalent to DMRG. The integration scheme, like DMRG, can be based either on one site or two sites. In the one-site method, the symplectic property is retained, but the bond dimension of the MPS cannot increase to accommodate increased entanglement. The two-site scheme involves a truncation process that allows evolution to a manifold with higher or lower bond dimension, although the symplectic property is lost.

A highly desirable feature of TEBD methods is that, aside from well-understood Trotter decomposition errors, the errors in the evolution all stem from the singular value decomposition (SVD) truncation of an MPS bond, which is precisely quantified and controlled. This is a much better situation than in ground-state DMRG, where the search for the ground state may get stuck in a local minimum[1][59, 60, 42].

Unfortunately, TDVP methods can fail in an uncontrolled manner, in some ways similar to DMRG. This failure in a simple situation is illustrated in Fig. 3.1(c). The system is a

[1]In DMRG, we can alleviate the problem using density matrix perturbation[59], subspace expansion[60], more center sites[42], or swap gates.

rung-decoupled Heisenberg ladder–when starting the time evolution in a product state, both the one- and two-site schemes fail to time-evolve at all.

This example illustrates some of the significant projection errors that can arise in TDVP. In the one-site scheme, TDVP evolution of a product state stays in a product state for any Hamiltonian. In addition, after being mapped to a one-dimensional geometry by using the zigzag (or snake) path, there exist two subsets of sites that are disconnected by interactions (Fig. 3.1(a)). The two sites involved in the local update at each bond (or each odd bond for the snake path) belong to the two disconnected subsets respectively, so two-site TDVP fails to build up the intra-leg entanglement unless the tangent space of the MPS manifold already contains the necessary degrees of freedom. When the initial state is a product state, i.e. an MPS of bond dimension one, the limitations of the tangent space are the most severe. In a more typical system, e.g. with the rung coupling turned on, as shown in Fig. 3.1(b), no disconnected subsets exist–although the two sites at even bonds are not directly connected by nearest-neighbor interactions, they are connected through a remote site. Therefore, in such situations the errors of two-site TDVP can be reduced by using a smaller time step.

There have been some tricks to enlarge the bond dimension of the product state to be time evolved by the one-site TDVP. The original MPS can be embedded in a MPS with larger bond dimension by filling up zeros[100, 101], but this approach does not help to immediately reduce the projection errors[63]. Another way is to use several DMRG sweeps to introduce some noise which artificially increases the bond dimension[50]. But our test shows that for two-dimensional systems and long-range interactions, large errors emerge after a short time even though the bond dimension has been enlarged substantially. Increasing the bond dimension to a large value at the beginning and keeping it through the whole time evolution is in fact very inefficient, since the bond dimension needed for the initial time might otherwise be much smaller than at later times. So it is suggested[102, 50] to first use two-site TDVP to increase the bond dimension for the initial sweeps and then switch to the one-site TDVP.

However, in the example in Fig. 3.1(c), two-site TDVP also fails.

In this paper we provide an effective way to fix the projection errors in TDVP methods, which works for both the extreme case above and other systems with long-range interactions even if a large time step is used. We show how one can expand the MPS manifold thereby enlarging the tangent space before a TDVP time step, to make it contain the true direction of motion. This enlargement is based on a subspace expansion[103, 60] by global Krylov vectors. Unlike in the exact diagonalization[104, 105], in our algorithm the global Krylov vectors serve as ancillary MPS's to enrich the basis of the time-evolving MPS through the gauge degree of freedom, thus avoiding the problems of loss of orthogonality and production of unnecessarily highly entangled state[50]. Our method has a modest computational cost compared to that of a TDVP time step. Also, it is easy to turn off the tangent space enlargement in regimes where ordinary TDVP methods work well.

3.2 Algorithms

Normally, in working with MPS, we always seek the smallest MPS to represent a particular state. Here, we find that it is very useful to temporarily create an inefficient MPS representation of the current time-evolved state by expanding the MPS to represent both the current state and a short Krylov expansion of it. The Krylov expansion is generated with a standard MPS algorithm[53, 50] which is capable of producing nonlocal entanglement. With the expanded manifold coming from this inefficient representation, a subsequent TDVP time step is accurate and reliable.

In this section, we first introduce the basis extension of an MPS by another one and discuss the trick to exactly preserve the information of the original time-evolved state. Then we discuss several issues of generating the global Krylov vectors. The algorithm is summarized

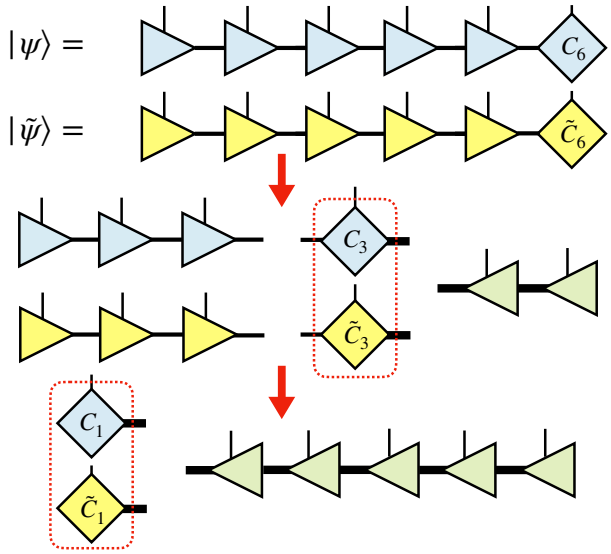

Figure 3.2: Basis extension of $|\psi\rangle$ by $|\tilde{\psi}\rangle$. The diamond $C_i(\tilde{C}_i)$ is the orthogonality center of $|\psi\rangle(|\tilde{\psi}\rangle)$ at site i. The triangle pointing to the right(left) is left(right)-orthonormal. The red dashed rounded rectangle means forming a direct sum of the tensors inside it. The bolder bond has the larger bond dimension.

in section 3.2.3.

3.2.1 Basis extension

The MPS representation of a physical state is not unique. We can utilize the gauge degrees of freedom to extend the basis at each bond so as to get an MPS with enlarged bond dimension without changing the physical state. Specifically, this property is used in our method to yield an MPS with its basis extended by other MPS's, which is reminiscent of the multi-state targeting[96, 106] approach frequently used in the early days of DMRG to deal with excited states simultaneously with the ground state. In the old DMRG language, targetting more than one state (also called state-averaging) is like having a mixed state, and one averages density matrices at each DMRG step. In the modern MPS language, the state-averaging is done by creating an extra index to label the states involved. In the description of our algorithm, we incorporate both formulations, which are equivalent.

Suppose we have an MPS of a state $|\psi\rangle$, and we wish to extend the MPS bond basis of $|\psi\rangle$ by that of another state $|\tilde{\psi}\rangle$. Suppose both states are in left-canonical form (Fig. 3.2),

$$|\psi\rangle = \sum_{s_1 \cdots s_N} A_1^{s_1} \cdots A_{N-1}^{s_{N-1}} C_N^{s_N} |s_1 \cdots s_N\rangle, \tag{3.1}$$

and similarly for $|\tilde{\psi}\rangle$. Here $A_1^{s_1}$ is a $1 \times m_1$ matrix and $C_N^{s_N}$ is a $m_{N-1} \times 1$ matrix, where m_i is the bond dimension between site i and $i+1$. First we write the direct sum formally,

$$\begin{aligned} \begin{bmatrix} |\psi\rangle \\ |\tilde{\psi}\rangle \end{bmatrix} &= \sum_{s_1 \cdots s_N} A_1'^{s_1} \cdots A_{N-1}'^{s_{N-1}} C_N'^{s_N} |s_1 \cdots s_N\rangle \\ &= \sum_{s_1 \cdots s_N} \begin{bmatrix} A_1^{s_1} & 0 \\ 0 & \tilde{A}_1^{s_1} \end{bmatrix} \cdots \begin{bmatrix} A_{N-1}^{s_{N-1}} & 0 \\ 0 & \tilde{A}_{N-1}^{s_{N-1}} \end{bmatrix} \begin{bmatrix} C_N^{s_N} \\ \tilde{C}_N^{s_N} \end{bmatrix} |s_1 \cdots s_N\rangle, \end{aligned} \tag{3.2}$$

where the second line defines the primed matrices in the first line. Now $A_1'^{s_1}$ is $2 \times (m_1 + \tilde{m}_1)$, and the product of matrices gives a 2×1 coefficient matrix (for each set of $s_1 \cdots s_N$). The extra two-dimensional index, attached to the first site, picks either $|\psi\rangle$ or $|\tilde{\psi}\rangle$. We can now compress the expanded MPS by doing SVD with truncation iteratively from the right end to the left. At site i, we perform $C_i'^{s_i} = U_i' S_i' B_i'^{s_i}$. We continue at the next site $i-1$ with

$$C_{i-1}' = \begin{bmatrix} C_{i-1} \\ \tilde{C}_{i-1} \end{bmatrix} = \begin{bmatrix} A_{i-1} C_i B_i'^{\dagger} \\ \tilde{A}_{i-1} \tilde{C}_i B_i'^{\dagger} \end{bmatrix}, \tag{3.3}$$

where the local physical indices have been omitted. Note that it is unnecessary to explicitly implement the form in Eq. (3.2). Because of the block diagonal form of A_{i-1}', it will be more efficient to move the orthogonality center to site $i-1$ separately for $|\psi\rangle$ and $|\tilde{\psi}\rangle$ as in Eq.

(3.3) than simply doing $C'_{i-1} = A'_{i-1}U'_iS'_i$. Eventually we will end up with

$$\begin{bmatrix} |\psi\rangle \\ |\tilde{\psi}\rangle \end{bmatrix} = \sum_{s_1\cdots s_N} \begin{bmatrix} C_1^{s_1} \\ \tilde{C}_1^{s_1} \end{bmatrix} B_2'^{s_2}\cdots B_N'^{s_N}|s_1\cdots s_N\rangle. \tag{3.4}$$

Note that this common representation for both states can be used to sum them, with arbitrary coefficients, $a|\psi\rangle + b|\tilde{\psi}\rangle$, by performing the same operation on $C_1^{s_1}$ and $\tilde{C}_1^{s_1}$. (One would then want to reorthogonalize from left to right, if one were only interested in the sum.) In our case, we want only $|\psi\rangle$, so we simply throw out $\tilde{C}_1^{s_1}$, which gives us a right-canonical MPS of $|\psi\rangle$ with its bond basis extended by $|\tilde{\psi}\rangle$ and its orthogonality center C_1 not full column rank.

However, the algorithm in this form has a drawback–it is not convenient to treat $|\tilde{\psi}\rangle$ less accurately than $|\psi\rangle$, or, in particular, to retain $|\psi\rangle$ exactly. To solve this problem, at each site i, instead of simply SVDing and truncating C'_i, we select the most important basis from $|\tilde{\psi}\rangle$ and orthogonalize them against the existing basis of $|\psi\rangle$ before combining them.

In the following we explain the reformed algorithm in more generic settings, with more than just two states—suppose we have k MPS's now and want to use the latter $k-1$ MPS's, $|\tilde{\psi}_l\rangle$, to extend the basis of the first one, $|\psi\rangle$. The extra time complexity to deal with k MPS's will be larger by $O(k)$. At site i (suppressing the index i below), we first SVD the site tensor of $|\psi\rangle$ without truncation, i.e. $C = USB$, and form the null-space projection operator $P = 1 - B^\dagger B$. We sum the reduced density matrices of the other $k-1$ MPS's

$$\tilde{\rho} = \sum_{l=1}^{k-1} \tilde{\rho}_l, \tag{3.5}$$

where $\tilde{\rho}_l = \tilde{C}_l^\dagger \tilde{C}_l$. If $P \neq 0$, we then project $\tilde{\rho}$ by P,

$$\bar{\rho} \equiv P\tilde{\rho}P. \tag{3.6}$$

Diagonalizing and truncating, $\bar{\rho} = \bar{B}^\dagger \bar{S}^2 \bar{B}$. (This is equivalent to projecting each $\tilde{C}_l$ by P and SVDing the direct sum of them.) The rows of $\bar{B}$ are orthogonal to those of B, i.e. $\bar{B}B^\dagger = 0$, so we can enlarge the row space of B by the direct sum

$$B' = \begin{bmatrix} B \\ \bar{B} \end{bmatrix}, \tag{3.7}$$

forming the new right-orthonormal MPS tensor at site i in Eq. (3.4).

3.2.2 Krylov subspace

What states do we use to enlarge the basis for $|\psi\rangle$? It is natural to consider the time evolution which TDVP is implementing. Consider the wavefunction evolved for a short time:

$$|\psi(t+\Delta t)\rangle = \exp\left(-\mathrm{i}\hat{H}\Delta t\right)|\psi(t)\rangle \approx \sum_{l=0}^{k-1} \frac{(-i\Delta t)^l}{l!}\hat{H}^l|\psi(t)\rangle, \tag{3.8}$$

where $t+\Delta t$ can be either imaginary or real. A problem with utilizing Eq. (3.8) directly for the time evolution is that the expression converges slowly in k. Instead, we use only a few of the terms appearing in Eq. (3.8) to extend the basis set of $|\psi(t)\rangle$ before time evolution by a following TDVP sweep. Specifically, in the algorithm presented here we consider an MPS extended to represent the Krylov subspace of order k

$$\mathcal{K}_k(\hat{H}, |\psi\rangle) = \mathrm{span}\{|\psi\rangle, \hat{H}|\psi\rangle \ldots, \hat{H}^{k-1}|\psi\rangle\}, \tag{3.9}$$

where the order k of the Krylov subspace is quite small.

There are three technical issues which need further elaboration. First, since the norm of $\hat{H}^l|\psi(t)\rangle$ grows exponentially with l, for numerical stability, we either normalize each MPS separately, or, motivated by the first-order expansion of the time evolution operator, replace

them by

$$(1 - \mathrm{i}\tau\hat{H})|\psi(t)\rangle, \ldots, (1 - \mathrm{i}\tau\hat{H})^{k-1}|\psi(t)\rangle, \tag{3.10}$$

where τ is a small parameter to be tuned to make sure the norm of $(1 - \mathrm{i}\tau\hat{H})^l|\psi(t)\rangle$ do not blow up. For imaginary time evolution, we can choose $\mathrm{i}\tau$ to be λ^{-1}, where λ is approximately the highest energy of the excited states. For real time evolution, the choice of τ does not matter as much and we can simply set $\tau = \Delta t$. Note the states in Eq. (3.10) still span the Krylov subspace.

The second issue is, how do we apply $\hat{H}$ efficiently? When the bond dimension of the MPO of $\hat{H}$ is small, we can use the density matrix approach[53] (which is exact if not truncated); otherwise we can use the variational approach[50]. The complexity of applying $\hat{H}$ at each site is comparable to one Lanczos iteration used to integrate the local effective equations at a site in TDVP. Usually the number of iterations needed at a site in TDVP is much larger than k, so the time cost of the application of $\hat{H}$ is subleading.

The third issue is, how do we control the bond dimension of $\hat{H}^l|\psi(t)\rangle$, which grows fast with increasing l? Fortunately, we have found that for a reasonable choice of time step size, $k = 3$ can already provide good accuracy. Furthermore, we do not need as high an accuracy in $\hat{H}^l|\psi(t)\rangle$ as compared to $|\psi(t)\rangle$ itself. Therefore we utilize a considerable truncation cutoff ϵ_K in applying $\hat{H}^l$.

There are a few other places in our algorithm where truncation is necessary. When diagonalizing the projected *mixing* of reduced density matrices in Eq. (3.6), we use a truncation cutoff ϵ_M. This truncation controls the number of states added into $|\psi(t)\rangle$. We also define a truncation cutoff ϵ which is applied during the TDVP time step. This cutoff is only used when we do not require the time evolution to be exactly unitary, such as in imaginary time evolution.

3.2.3 Subspace expansion

We now describe our *global subspace expansion* (GSE) TDVP algorithm. Because of the Krylov subspace expansion, we do not need to use a two-site method. Combined with the one-site TDVP, one time step of our algorithm (GSE-TDVP1) consists of

1. Construct the MPO for $1 - \mathrm{i}\tau\hat{H}$, and apply it $k-1$ times to $|\psi(t)\rangle$ and get the set of MPS's in Eq. (3.10). (Controlled by ϵ_K.)

2. Do basis extension for $|\psi(t)\rangle$ as described in 3.2.1, ending with the orthogonality center at the first site. (Controlled by ϵ_M.)

3. From left to right, do a conventional one-site TDVP sweep, and then sweep from right to left. (Controlled by ϵ.)

Let the bond dimension of $|\psi(t)\rangle$ before step 1 be m and the bond dimension of the MPO for $1 - \mathrm{i}\tau\hat{H}$ be w. Usually we choose ϵ_K to make the bond dimension of each of the Krylov vectors in Eq. (3.10) no bigger than m. Then, if we apply the MPO variationally[50], the complexity of step 1 is $O(km^3wd)$. Let the bond dimension of $|\psi(t)\rangle$ after step 2 be m'. Then the complexity of step 2 is $O(km'^3d^2 + m'^3d^3)$. The one-site TDVP has a complexity $O(lm'^3wd)$, where l is the number of Lanczos steps used at each local update. As described in the last section, usually $l \gg k$ so the cost of step 1 and 2 will be comparable to or less than step 3. In our benchmarks for real time evolution, we find that if we use $k = 5$ and get $m' \approx 3m$, the time taken for step 1 and 2 is about 1/3 of that for step 3; given the same bond dimension m', the time taken for one GSE-TDVP1 time step is about 36% of that for a conventional two-site TDVP step.

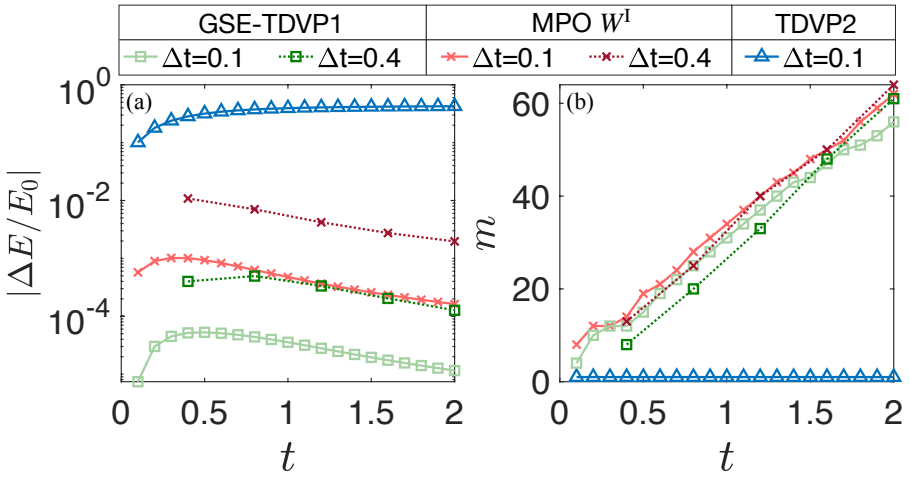

Figure 3.3: Benchmark results of the imaginary time evolution for the rung-decoupled Heisenberg ladder. In all methods, $\epsilon = 10^{-10}$. For MPO W^{I}, ϵ is the truncation error in applying the MPO. For GSE-TDVP1, we use the optimal settings i$\tau = 1/40$, $k = 3$, $\epsilon_M = 10^{-8}$, and ϵ is the truncation error in the follow-up TDVP1 sweep. (a) Absolute energy error $\Delta E = E - E_{\mathrm{exact}}$ scaled by the ground-state energy E_0, where the reference energy E_{exact} is obtained by doubling the energy of a single chain from TDVP2 with $\Delta t = 0.01$. (b) Bond dimension growth versus time. All methods except TDVP2 (which stays in a product state) show similar bond dimension growth.

3.3 Benchmarks

3.3.1 Imaginary time evolution

We first consider imaginary time evolution of the rung-decoupled spin-1/2 Heisenberg ladder of Fig. 3.1, with Hamiltonian

$$\hat{H} = \sum_{r,\langle i,j\rangle} \hat{\boldsymbol{S}}_{r,i} \cdot \hat{\boldsymbol{S}}_{r,j}, \tag{3.11}$$

where $r \in \{1, 2\}$ denotes which leg it is and $\langle i, j\rangle$ denotes the nearest-neighbor sites along each leg.

We evolve in imaginary time starting from a Neel (product) state $|\psi(0)\rangle$. Measuring the energy versus time provides a reasonable test of the evolution. In Fig. 3.3 we show a

comparison of our method with two other methods. The first one is an MPO method of Ref. [17], specifically the W^{I} method, where complex time steps have been used to make the error second-order. The second is the conventional TDVP2. In our method, we use $\epsilon_K = 10^{-12}$ when applying $1 - \mathrm{i}\tau\hat{H}$ with the density matrix approach[2]. We show results for $\mathrm{i}\tau = 1/40$, $k = 3$, and $\epsilon_M = 10^{-8}$, which turns out to be near optimal parameter settings in this case. We found that higher order k or smaller truncation ϵ_M in the subspace expansion do not improve the accuracy (not shown). While TDVP2 fails as expected, GSE-TDVP1 (with $\epsilon = 10^{-10}$) has an accuracy 10 times better than the MPO W^{I} method with a comparable bond dimension growth. We also show GSE results with a larger time step $\Delta t = 0.4$, which exhibit a still reasonable error of 10^{-3}. Not shown are results for $\epsilon = 10^{-12}$, which are slightly more accurate than $\epsilon = 10^{-10}$, but which also exhibit faster bond dimension growth.

3.3.2 Real time evolution

The one-axis twisting (OAT) model[107] has been widely studied for the use of quantum metrology[108]. This model has infinite-range interactions, making it a challenge for MPS methods, but it also has an exact solution. We study the real time evolution of $N = 100$ spin-1/2's with the Hamiltonian

$$\hat{H} = \chi(\hat{S}^z)^2, \tag{3.12}$$

where $\hat{S}^z = \sum_{i=1}^{N} \hat{S}_i^z$ and we set the energy scale $\chi = 1$. The MPO representation of the Hamiltonian is rather simple, with a bond dimension $w = 3$. We take as initial state all spins polarized in the $+x$ direction. During the time evolution the spin is squeezed[109].

[2]The variational approach faces the same local minimum problem as DMRG does, which can be solved by ways suggested before.

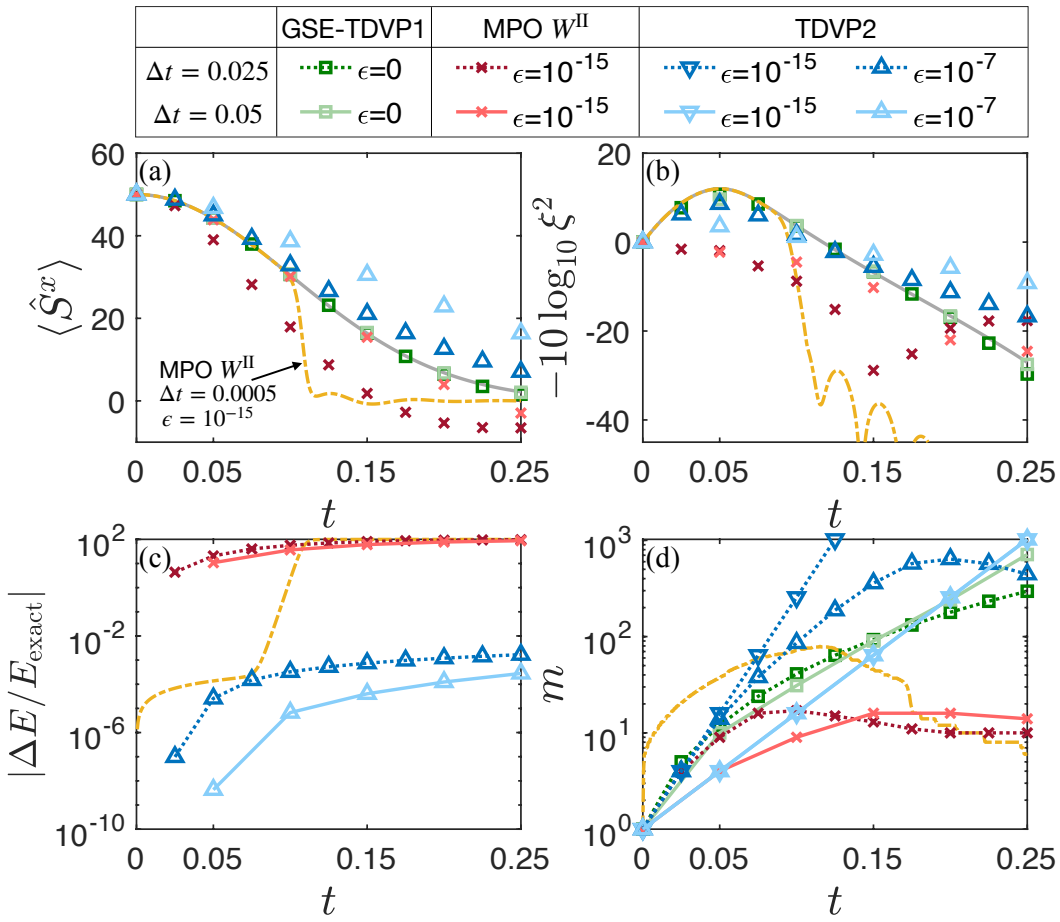

Figure 3.4: Benchmark results of the real time evolution for the OAT model for a variety of methods, versus exact results (solid grey lines). (a) shows $\langle \hat{S}^x(t) \rangle$. GSE-TDVP1 is the most accurate method. The apparent accuracy of one of the MPO W^{II} calculations appears to be accidental. TDVP2 becomes more accurate with a smaller time step. (b) shows the spin squeezing parameter. Again, GSE-TDVP1 is the most accurate. (c) shows relative errors in the energy. Curves for GSE-TDVP1 and for TDVP2 with $\epsilon = 10^{-15}$ are all below 10^{-10} and are not shown. Errors for the MPO W^{II} are particularly large. (d) shows the bond dimension growth. The smallest bond dimensions come from the MPO W^{II} methods but this is due to their large errors. TDVP2 with smaller time step has particularly large bond dimension growth. Note that in both (a) and (b) the TDVP2 data points for $\epsilon = 10^{-15}$ (not shown) and $\epsilon = 10^{-7}$ coincide.

The exact solution for the x-moment is

$$\langle \hat{S}^x(t) \rangle = \frac{N}{2} \cos^{N-1}(t), \tag{3.13}$$

with $\langle \hat{S}^y(t) \rangle = \langle \hat{S}^z(t) \rangle = 0$.

An important property is the spin squeezing parameter, defined[110] as

$$\xi^2 = N \min_{\mathbf{n}_\perp} \frac{\langle (\hat{\boldsymbol{S}} \cdot \mathbf{n}_\perp)^2 \rangle - \langle \hat{\boldsymbol{S}} \cdot \mathbf{n}_\perp \rangle^2}{\langle \hat{\boldsymbol{S}} \rangle^2}, \tag{3.14}$$

where $\hat{S}^\mu = \sum_i \hat{S}_i^\mu$ and $\mathbf{n}_\perp$ is a unit vector perpendicular to $\langle \hat{\boldsymbol{S}} \rangle = \langle \hat{S}^x \rangle \mathbf{n}_x$. Minimizing over $\mathbf{n}_\perp$, ξ^2 can be expressed in terms of correlation functions

$$\frac{\xi^2}{N} = \frac{\sigma_{yy}^2 + \sigma_{zz}^2 - \sqrt{(\sigma_{yy}^2 - \sigma_{zz}^2)^2 + (\sigma_{yz}^2 + \sigma_{zy}^2)^2}}{2\langle \hat{S}^x \rangle^2}, \tag{3.15}$$

where $\sigma_{xy}^2 \equiv \langle \hat{S}^x(t) \hat{S}^y(t) \rangle$, etc. The optimal spin squeezing ξ_{opt}^2 is expected to appear at $t_{\text{opt}} = 12^{\frac{1}{6}} (N/2)^{-\frac{2}{3}}/2 \approx 0.05$. We continue the time evolution to $t = 0.25$, or about $5t_{\text{opt}}$.

In Fig. 3.4, we compare our method with TDVP2 and MPO W^{II} (which is expected to work better than W^{I} here). To preserve exact unitarity, for GSE-TDVP1 we set $\epsilon = 0$, which turns out to have a minimal extra cost in bond dimension, with the bond dimension already controlled by ϵ_K and ϵ_M. Since w is small, we use the density matrix approach to apply the MPO. For time step $\Delta t = 0.025$, we find optimal parameters $\tau = \Delta t$, $k = 3$, $\epsilon_K = 10^{-4}$, and $\epsilon_M = 10^{-4}$, which balance cost and accuracy. For time step $\Delta t = 0.05$, we use $\tau = \Delta t$, $k = 5$, $\epsilon_K = 10^{-4}$, and $\epsilon_M = 10^{-8}$. Our method is the most accurate and preserves unitarity exactly, while also having slower bond dimension growth than TDVP2. For MPO W^{II}, the conservation of energy is very poor and the overall shape of ξ^2 is wrong. Reducing the time step size to $\Delta t = 0.0005$ (yellow curves in Fig. 3.4) helps initially, but the evolution soon becomes unstable after $t_{\text{opt}} \approx 0.05$.

Chapter 4

Spin squeezing dynamics on two-dimensionl long-range XXZ model

This work to be published is a collaboration with Sean R. Muleady and Prof. Ana Maria Rey at University of Colorado, Boulder.

4.1 Introduction

Quantum spin squeezing[108] is useful in detecting quantum entanglement and improving the precision of measurements such as in Ramsey spectroscopy[110], atomic clocks[111], and gravitational-wave interferometers[112, 113].

4.1.1 Coherent spin state

Let us first review how coherent states are defined for bosons and then go to the spins.

The Heisenberg uncertainty relation gives

$$\Delta X \Delta P \geq 1, \tag{4.1}$$

where $\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}$ is the standard deviation, and

$$X \equiv a^\dagger + a, \quad P \equiv \mathrm{i}(a^\dagger - a) \tag{4.2}$$

with a and $a^\dagger$ being the bosonic creation and annihilation operators that satisfy

$$[a, a^\dagger] = 1. \tag{4.3}$$

A coherent state defined as

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{4.4}$$

is the minimum-uncertainty state that obeys

$$\Delta X = \Delta P = 1. \tag{4.5}$$

For example, the vacuum state $|0\rangle$ is a coherent state.

Now we switch to spin systems. In the field of ultra cold atoms, a two-level atom interacting with a radiation field can be treated as a spin-1/2 particle in a magnetic field. The coherent spin state (CSS) for N spin-1/2 particles is defined as a product of single spin states with

all spins point to the same direction

$$
\begin{aligned}
|\theta,\phi\rangle &= \bigotimes_{l=1}^{N}\Big[R_l(\theta,\phi)|\uparrow\rangle_l\Big] \\
&= \bigotimes_{l=1}^{N}\Big[\exp\left(-\mathrm{i}\theta\sigma^l\cdot\mathbf{n}/2\right)|\uparrow\rangle_l\Big] \\
&= \bigotimes_{l=1}^{N}\Big[\cos\frac{\theta}{2}|\uparrow\rangle_l + e^{\mathrm{i}\phi}\sin\frac{\theta}{2}|\downarrow\rangle_l\Big],
\end{aligned}
\tag{4.6}
$$

where $\sigma^l = (\sigma_x^l, \sigma_y^l, \sigma_z^l)$ are the Pauli matrices for the lth particle and $\mathbf{n} = (-\sin\phi, \cos\phi, 0)$. We can use the angular momentum operator of N spin-1/2 particles

$$
J_\alpha = \frac{1}{2}\sum_{l=1}^{N}\sigma_\alpha^l, \qquad \alpha = x, y, z \tag{4.7}
$$

to further write the CSS as

$$
|\theta,\phi\rangle = R(\theta,\phi)|j,j\rangle = \exp\left(-\mathrm{i}\theta\mathbf{J}\cdot\mathbf{n}\right)|j,j\rangle, \tag{4.8}
$$

where $|j,j\rangle = \bigotimes_{l=1}^{N}|\uparrow\rangle_l$ is the eigenstate of J_z with eigenvalue $j = N/2$. We can express the CSS in terms of the Dicke states $|j,m\rangle$

$$
|\theta,\phi\rangle = \sum_{m=-j}^{j}|j,m\rangle\langle j,m|\theta,\phi\rangle = (1+|\eta|^2)^{-j}\sum_{m=-j}^{j}\begin{pmatrix} 2j \\ j+m \end{pmatrix}^{1/2}\eta^{j+m}|j,m\rangle, \tag{4.9}
$$

where $\eta = -\tan\frac{\theta}{2}\exp\left(-\mathrm{i}\phi\right)$. When $\theta = \pi/2$, $|\langle j,m|\theta,\phi\rangle|^2$ is called projection noise which obeys a binomial distribution (approximating to Gaussian when $j \to \infty$ according to the central limit theorem).

4.1.2 Spin squeezing

We first review the basics of bosonic squeezed states and then show how the spin squeezing can be defined analogously.

A bosonic state which makes ΔX (or ΔP) smaller than one is called a squeezed state. The operator

$$X_\theta = e^{\mathrm{i}\theta a^\dagger a} X e^{\mathrm{i}\theta a^\dagger a} = a e^{-\mathrm{i}\theta} + a^\dagger e^{\mathrm{i}\theta} \tag{4.10}$$

is obtained by rotating X in the $X - P$ plane. $X_0 = X$ and $X_{\frac{\pi}{2}} = P$ are the special cases. The principal-quadrature squeezing parameter is defined as

$$\xi_B^2 = \min_\theta (\Delta X_\theta)^2 = 1 + 2(\langle a^\dagger a\rangle - |\langle a\rangle|^2) - 2|\langle a^2\rangle - \langle a\rangle^2|. \tag{4.11}$$

A bosonic state is squeezed if $\xi_B^2 < 1$. A bosonic squeezed state can be generated by the real time evolution of a coherent state with a nonlinear Hamiltonian

$$H = -\mathrm{i}(g a^{\dagger 2} - g^* a^2) \tag{4.12}$$

or the Kerr interaction[114]

$$H = \kappa (a^\dagger a)^2. \tag{4.13}$$

There are several definitions of spin squeezing. One definition ξ_S^2 was inspired by photon squeezing[107]. Another well-known definition ξ_R^2 was proposed naturally in standard Ramsey spectroscopy[110], which directly associated with quantum metrology.

Consider the uncertainty relation for angular momentum operators

$$(\Delta J_\alpha)^2(\Delta J_\beta)^2 \geq \frac{|\langle J_\gamma\rangle|^2}{4}, \tag{4.14}$$

where α, β, and γ denotes the components in any three orthogonal directions. In analogy to bosonic squeezed states, a spin squeezed state can be defined as a state that makes one component of the angular momentum satisfy

$$(\Delta J_\alpha)^2 < \frac{|\langle J_\gamma\rangle|^2}{2}, \tag{4.15}$$

and the corresponding spin squeezing parameter will be

$$\xi_H^2 = \frac{2(\Delta J_\alpha)^2}{|\langle J_\gamma\rangle|^2}, \quad \alpha\perp\gamma. \tag{4.16}$$

However in this definition a CSS can also have $\xi_H^2 < 1$. So the definition need to be fixed. Different from bosonic systems, for a CSS $(\Delta J_\alpha)^2$ depends on the direction α. The spin squeezing parameter in Ref. [107] is defined as

$$\xi_S^2 = \frac{\min_{\mathbf{n}_\perp}(\Delta J_{\mathbf{n}_\perp}^2)}{j/2} = \frac{4\min_{\mathbf{n}_\perp}(\Delta J_{\mathbf{n}_\perp}^2)}{N}, \tag{4.17}$$

where $\mathbf{n}_\perp$ is the direction perpendicular to the mean spin direction (MSD)

$$\mathbf{n}_0 = \frac{\langle \mathbf{J}\rangle}{|\langle \mathbf{J}\rangle|} \tag{4.18}$$

so $\langle J_{\mathbf{n}_\perp}\rangle = 0$. Now for a CSS $\xi_S^2 = 1$ since $(\Delta J_{\mathbf{n}_\perp})^2 = j/2$ if the state is a CSS.

Ref. [110] defined another type of spin squeezing parameter which is the ratio of the phase

resolution between a correlated state and a CSS, i.e.

$$\xi_R^2 = \frac{(\Delta\phi)^2}{(\Delta\phi_{\mathrm{CSS}})^2}, \tag{4.19}$$

where $\Delta\phi$ is the phase resolution of the resonance frequency in Ramsey spectroscopy. $\xi_R^2 < 1$ characterizes the improvement of the sensitivity to a rotation angle ϕ for a spin squeezed state compared to a CSS. If we rotate a general spin state with MSD $\mathbf{n}_0 = z$ about the x-axis by angle ϕ, we will get

$$\begin{aligned} &\langle J_y\rangle' = -\sin\phi\langle J_z\rangle, \\ &(\Delta J_y')^2 = \cos^2\phi(\Delta J_y)^2 + \sin^2\phi(\Delta J_z)^2 - \frac{1}{2}\sin(2\phi)\langle\{J_y, J_z\}\rangle. \end{aligned} \tag{4.20}$$

The phase sensitivity $\Delta\phi$ is calculated by the error propagation

$$\Delta\phi = \frac{\Delta J_y'}{|\partial\langle J_y\rangle'/\partial\phi|} = \frac{\Delta J_y'}{|\cos\phi\langle J_z\rangle|}. \tag{4.21}$$

For tiny rotation angle ϕ, $\Delta J_y' \sim \Delta J_y$ and $\cos\phi \sim 1$, so we have

$$\Delta\phi = \frac{\Delta J_y}{|\langle J_z\rangle|}. \tag{4.22}$$

More generally when $\mathbf{n}_0 \neq z$, we have

$$\Delta\phi = \frac{\Delta J_{\mathbf{n}_\perp}}{|\langle\mathbf{J}\rangle|}, \tag{4.23}$$

so the spin squeezing parameter can be further written as

$$\xi_R^2 = \frac{N\min_{\mathbf{n}_\perp}(\Delta J_{\mathbf{n}_\perp})^2}{|\langle\mathbf{J}\rangle|^2}, \tag{4.24}$$

where we already use the standard quantum limit

$$\Delta\phi_{\mathrm{CSS}} = \frac{1}{\sqrt{N}}. \tag{4.25}$$

There is also a lower bound $\Delta\phi \geq 1/N$, called the Heisenberg limit, when the equality in the uncertainty relation is reached.

For a general spin state $|\psi\rangle$ with MSD $\mathbf{n}_0$, we can expand it in terms of the Dicke states in a $\mathbf{n}_\perp$ direction, i.e.

$$|\psi\rangle = \sum_{m=-j}^{j} |j,m\rangle_{\mathbf{n}_\perp\mathbf{n}_\perp}\langle j,m|\psi\rangle, \tag{4.26}$$

and we call $P(m) = |_{\mathbf{n}_\perp}\langle j,m|\psi\rangle|^2$ projection noise, which gives the probability of getting a angular momentum m in $\mathbf{n}_\perp$ direction in measurements. Since

$$\begin{aligned}
\langle J_{\mathbf{n}_\perp}\rangle &= \sum_{m=-j}^{j} \langle\psi|J_{\mathbf{n}_\perp}|j,m\rangle_{\mathbf{n}_\perp\mathbf{n}_\perp}\langle j,m|\psi\rangle = \sum_{m=-j}^{j} P(m)m, \\
\langle J^2_{\mathbf{n}_\perp}\rangle &= \sum_{m=-j}^{j} \langle\psi|J^2_{\mathbf{n}_\perp}|j,m\rangle_{\mathbf{n}_\perp\mathbf{n}_\perp}\langle j,m|\psi\rangle = \sum_{m=-j}^{j} P(m)m^2,
\end{aligned} \tag{4.27}$$

the variance $(\Delta J_{\mathbf{n}_\perp})^2$ or the phase sensitivity $\Delta\phi$ purely depends on the distribution of $P(m)$ along the direction $\mathbf{n}_\perp$. For a CSS, $P(m)$ is binomial and identical in any direction $\mathbf{n}_\perp$; for a coherent spin state, $P(m)$ might be sub-binomial[115] and there is a direction that the distribution has the minimal width.

The spin squeezing parameter is actually a demonstration of quantum correlations since it is smaller than one only when correlations (entanglement) are built among the spins, which results in a reduced fluctuation in certain direction.

4.1.3 One-axis twisting model

The one-axis twisting (OAT) model[107] was inspired by the nonlinear Hamiltonian to generate squeezed photon. In the following we show some relevant analytical results[108] about the spin squeezing dynamics under the one-axis twisting Hamiltonian.

The one-axis twisting Hamiltonian

$$H = \chi (J^z)^2, \tag{4.28}$$

is nonlinear and involves all pairwise interactions. The initial state is taken to be a CSS $|\psi_0\rangle$ with all N spins polarized to the $+x$ direction. Since

$$[H, P_x] = 0, \tag{4.29}$$

where P_x is the reflection operator about the $y-z$ plane, it can be easily proved that $\langle\psi_0|e^{\mathrm{i}Ht}J_y e^{-\mathrm{i}Ht}|\psi_0\rangle = \langle\psi_0|e^{\mathrm{i}Ht}J_z e^{-\mathrm{i}Ht}|\psi_0\rangle = 0$. So the MSD $\mathbf{n}_0 = x$ throughout the time evolution. Using the commutation relation for Pauli matrices and exchange symmetry of spins, we can get

$$\langle\psi_0|e^{\mathrm{i}Ht}J_x e^{-\mathrm{i}Ht}|\psi_0\rangle = \frac{N}{2}\cos^{N-1}(\chi t). \tag{4.30}$$

Minimizing over $\mathbf{n}_\perp$, the spin squeezing parameter ξ_R^2 becomes

$$\xi_R^2 = \frac{N\left(\sigma_{yy}^2 + \sigma_{zz}^2 - \sqrt{(\sigma_{yy}^2 - \sigma_{zz}^2)^2 + (\sigma_{yz}^2 + \sigma_{zy}^2)^2}\right)}{2\langle J^x\rangle^2} \tag{4.31}$$

with the minimization direction

$$\mathbf{n}_\perp = (0, \cos\theta, \sin\theta) \tag{4.32}$$

where

$$\tan\theta = \frac{\sigma_{yy}^2 - \sigma_{zz}^2}{\sigma_{yz}^2 + \sigma_{zy}^2} \tag{4.33}$$

and

$$\begin{aligned}
\sigma_{yz}^2 + \sigma_{zy}^2 &\equiv \langle\psi_0|e^{\mathrm{i}Ht}(J_yJ_z + J_zJ_y)e^{-\mathrm{i}Ht}|\psi_0\rangle = \frac{N(N-1)}{2}\sin(\chi t)\cos^{N-2}(\chi t), \\
\sigma_{yy}^2 &\equiv \langle\psi_0|e^{\mathrm{i}Ht}J_yJ_ye^{-\mathrm{i}Ht}|\psi_0\rangle = \frac{N}{4} + \frac{N(N-1)}{8}\left(1 - \cos^{N-2}(2\chi t)\right), \\
\sigma_{zz}^2 &\equiv \langle\psi_0|e^{\mathrm{i}Ht}J_zJ_ze^{-\mathrm{i}Ht}|\psi_0\rangle = \frac{N}{4}.
\end{aligned} \tag{4.34}$$

The optimal squeezing happens at

$$t_{opt} = \frac{3^{\frac{1}{6}}}{N^{\frac{2}{3}}\chi}. \tag{4.35}$$

The amount of squeezing scales with system size, which is a desirable property of the OAT model. When N is very large, $\xi_R^2 \sim N^{-2/3}$ and $\theta \sim 1/2\arctan(N^{-1/3})$ at t_{opt}.

4.2 XXZ model with $1/r^\alpha$ interaction

Although the OAT model can be realized experimentally in a two-component Bose-Einstein condensation by using large-detuned light-atom interactions[108], it is still quite difficult to realize such all-to-all interactions in practice[109].

People have considered[116] the Ising model with power-law decaying interactions

$$H_{\text{Ising}} = \sum_{i\neq j}\frac{J}{|\mathbf{r}_i - \mathbf{r}_j|^\alpha}s_{z,i}s_{z,j}. \tag{4.36}$$

However, the amount of squeezing only scales with system size when $\alpha < D$ and is indepen-

dent of system size when $\alpha > D$, where D is the space dimension of the system. We still face the experimental difficulty to realize $\alpha < D$ interactions.

Researches[117, 118, 119] have shown that collective behavior like that in OAT can be energetically protected by an additional spin-exchange term in the Ising model. Then it is natural to ask the question if scalable spin squeezing can be achieved in this variant with $\alpha > D$ interactions.

Our collaborators proposed a XXZ model with power-law decaying interactions

$$H_{\mathrm{XXZ}} = \sum_{i \neq j} \frac{J_\perp \mathbf{s}_i \cdot \mathbf{s}_j + (J_z - J_\perp) s_{z,i} s_{z,j}}{|\mathbf{r}_i - \mathbf{r}_j|^\alpha}, \tag{4.37}$$

with the hope to interpolate between the OAT and the Ising model. When $\alpha \leq D$ and near the Heisenberg point $J_z = J_\perp$, the Dicke manifold is gapped away from all other states by an energy difference $\Delta \gtrsim |J_\perp|$, which allows for a perturbative treatment of the Ising term[109]. When $\alpha > D$, $\Delta \to 0$ as $N \to \infty$, perturbation theory becomes invalid, so our collaborators[109] numerically explored this regime by discrete truncated Wigner approximation (DTWA)[120, 121, 122] which can reach system size of $O(10^3)$. However, the validity of this classical method is remained to check. So we use our GSE-TDVP method to study the same model and compare with their results.

4.3 Discrete truncated Wigner approximation

In this section, we give a short introduction to the DTWA method.

Wigner function[122] is a one-to-one mapping between the density matrix operator to an ordinary function defined in the phase space. The truncated Wigner approximation (TWA) replaces the quantum-mechanical time evolution by a semi-classical evolution via classical

trajectories. The quantum uncertainty in the initial state is accounted for by an average over different initial conditions determined by a continuous Wigner function. DTWA[120] is a semiclassical method which is based on the Monte Carlo sampling of a discrete Wigner function in the phase space, which involves a factorization of the density matrix as $\hat{\rho} \approx \otimes_i \hat{\rho}_i$, where $\hat{\rho}_i$ is a local density matrix at lattice site i. The use of discrete Wigner functions enables us to quantitatively access dynamics in generic spin lattice models, including oscillations and revivals of single particle observables and correlation functions that are not captured by TWA.

In mean-field theory, the correlators are approximated by their mean-field values

$$\langle \hat{S}_i^\alpha \hat{S}_j^\beta \rangle \approx s_i^\alpha s_j^\beta, \qquad s_i^\alpha = \langle \hat{S}_i^\alpha \rangle. \tag{4.38}$$

If the Hamiltonian has the form

$$H = \sum_{i \neq j, \alpha\beta} \hat{S}_i^\alpha \hat{S}_j^\beta, \tag{4.39}$$

we will arrive at a set of classical equations of motion

$$\dot{s}_i^\alpha = \sum_{j \neq i, \beta\gamma} C_{ij}^{\alpha\beta\gamma} s_i^\beta s_j^\gamma, \tag{4.40}$$

where

$$C_{ij}^{\alpha\beta\gamma} = -\sum_\nu \epsilon^{\alpha\beta\nu} (h_{ij}^{\nu\gamma} + h_{ji}^{\gamma\nu}). \tag{4.41}$$

DTWA also involves solving such classical equation of motions. Nevertheless, different from mean-field theory, in DTWA the initial values of s_i^α are set to its classical values according to a probability distribution that reproduces the correct $\langle \hat{S}_i^\alpha \rangle$ at the initial time. So it is able to build up non-trivial covariances between the spins, i.e. $\langle \hat{S}_i^\alpha \hat{S}_j^\beta \rangle \approx \overline{s_i^\alpha s_j^\beta} \neq \bar{s}_i^\alpha \bar{s}_j^\beta$, where

the overline denotes the statistical average over different classical trajectories.

4.4 Comparison results

Our collaborator's DTWA simulation[109] on various system sizes up to 64×64 showed that collective behavior and thus considerable amount of spin squeezing can still happen in a large parameter regime when $\alpha > 2$. The optimal spin squeezing parameter in this regime has the favorable scaling with system size.

Since the system has long-range interactions, we use our GSE-TDVP1 method with $k = 3$, $\epsilon_K = 10^{-8}$, and $\epsilon_M = 10^{-4}$ for the initial 5 sweeps and then switch to the two-site TDVP method with a truncation error 10^{-8} for the following sweeps. Since t_{opt} can be qualitatively regarded as a measure of the speed of entanglement growth. We fix the number of sweeps from $t = 0$ to $t = t_{opt}$ to be 24 and the time step size is determined accordingly. By comparing with GSE-TDVP1 simulations with larger k, we can estimate the error in the spin squeezing parameter ξ_R^2 is about 10^{-2}. We simulate on system sizes up to 10×10. FIG. 4.1 shows various comparisons between GSE-TDVP and DTWA. We can see an improving agreement between the two different methods, which indicates the effectiveness of the DTWA on even larger systems.

An interesting question is about the steady state the spin dynamics approaches after long time. A generalized Mermin-Wagner theorem[123] allows for the existence of long-range order in the thermodynamic limit below the critical temperature when $\alpha < 4$ in two dimension. And only when $\alpha > 2$ is the energy an extensive quantity. So in the regime $2 < \alpha < 4$ we may expect the system to thermalize[124] to an equilibrium ordered phase in the long-time limit as long as the corresponding temperature is below the critical value. If we regard $\langle \hat{S}_{tot}^2 \rangle$ as an order parameter, FIG. 2 in Ref. [109] will be reminiscent of the zero phase diagram[125]

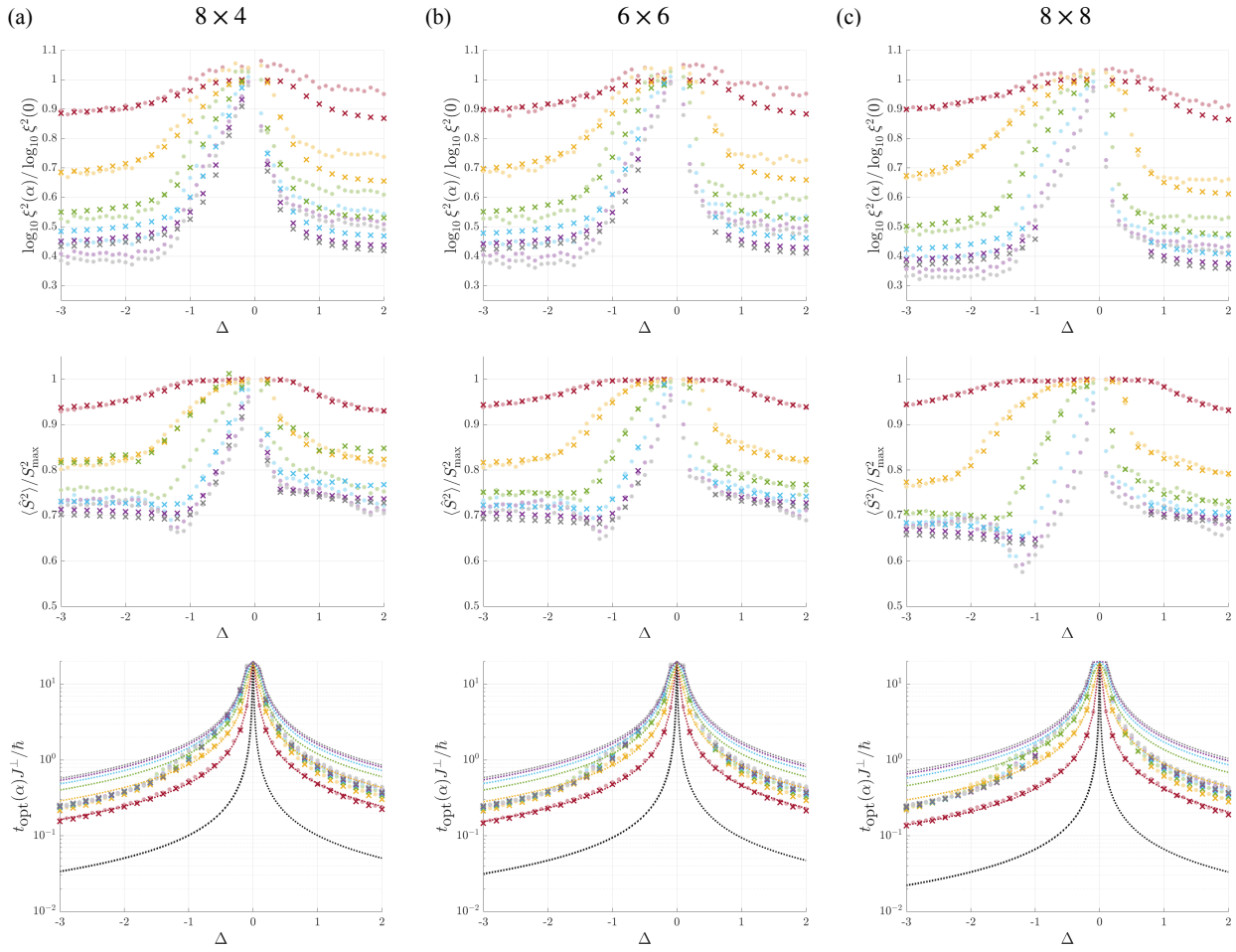

Figure 4.1: Comparisons of the spin squeezing parameter ξ_R^2, the expectation value of $\langle \hat{S}_{tot}^2 \rangle$ at t_{opt}, and t_{opt} between GSE-TDVP and DTWA for $\Delta = (J_z - J_\perp)/J_\perp$ from -3 to 2 and system sizes (a) 8×4, (b) 6×6, and (c) 8×8. The cross symbols are data points from GSE-TDVP and the circles are from DTWA. The colors follow the convention in FIG. 4.2.

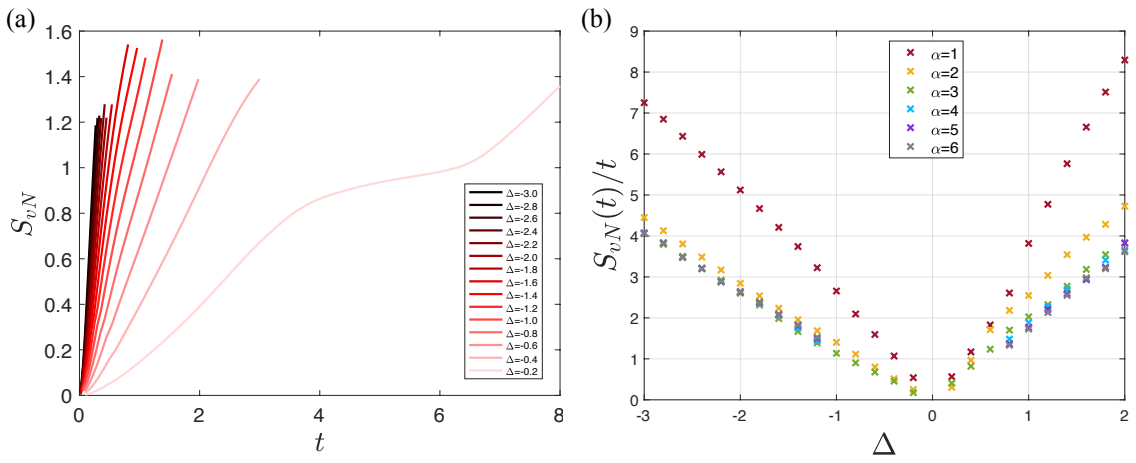

Figure 4.2: (a) The von Neumann entanglement entropy growth for a 8×8 system with $\alpha = 3$ calculated by GSE-TDVP. (b) The coefficient extracted from the linear fitting of $S_{vN}(t)$ for the same system.

of the same H_{XXZ} with $J_\perp < 0$.

The entanglement entropy calculated by our GSE-TDVP method shows an almost linear growth after a short time (FIG. 4.2), which is an indication of thermalization. Assuming that the system thermalizes, the equilibrium ensemble it thermalizes to should be unique. So we only to obtain the thermodynamic observables in the corresponding equilibrium ensemble and see if they match the values given by the spin squeezing dynamics. For a 6×6 system, we use our GSE-TDVP method combined with purification to imaginary time evolve the mixed state at infinite temperature to lower temperatures until its energy matches the conserved energy in the real time dynamics, calculating $(\hat{S}^x_{tot})^2 + (\hat{S}^y_{tot})^2$ and $(\hat{S}^z_{tot})^2$, which can be understood as the XY order parameter and the ferromagnetic order parameter respectively. From FIG. 4.3(b), we can see that when $\Delta < 0$ the thermal average of $(\hat{S}^z_{tot})^2$ match pretty well with its conserved real time value $N/4 = 9$, but they deviate from each other when $\Delta > 0$, which indicates we use the wrong ensemble in this case. Since $[(\hat{S}^z_{tot})^2, H] = 0$, we should impose en extra Lagrange multiplier in the Hamiltonian to take the conserved operator into account. For $\Delta > 0$, long-range ferromagnetic order might exist[1], so it would

[1] $\langle \hat{S}^z_{tot} \rangle$ is approximately zero all Δ in the imaginary time evolution, since we do not apply a pinning field, even when $\Delta > 0$ it might spontaneously symmetry breaks to a ferromagnetic phase.

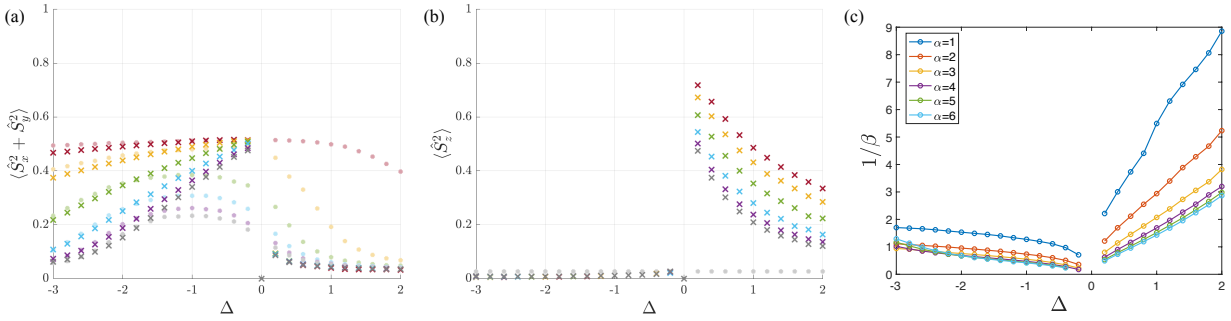

Figure 4.3: Comparison of the XY order parameter $(\hat{S}^x_{tot})^2 + (\hat{S}^y_{tot})^2$ and ferromagnetic order parameter $(\hat{S}^z_{tot})^2$ between the long-time limit of the spin squeezing dynamics and the thermal average of a canonical ensemble at the corresponding temperature $1/\beta$ for a 6×6 system. The cross symbols are data points from GSE-TDVP calculation on the equilibrium ensemble and the circles are from DTWA for steady state of the spin squeezing dynamics. The colors in (a) and (b) follow the convention in FIG. 4.2.

be a problem if the extra $\lambda(\hat{S}^z_{tot})^2$ were not used. It is not expected that the thermal averages of $(\hat{S}^x_{tot})^2 + (\hat{S}^y_{tot})^2$ do not match their long-time limit values when $-3/2 \lesssim \Delta < 0$ for $\alpha \geq 3$ (FIG. 4.3(a)). The system might get stuck in a prethermal state. Notice the difference between $2 < \alpha < 4$ and other alpha's only show up in the thermodynamic limit, which should be much larger than 6×6.

Chapter 5

Conclusion

We developed a DMRG approach to study continuum multi-species systems in one dimension, interacting with non-local Coulomb-like potentials. In order to get good convergence with the number of sweeps, we implemented a three-site DMRG algorithm, which performs well. As a first application, we have applied it to a model of 1D diatomic molecules, where we consider effects beyond the Born-Oppenheimer approximation. The most interesting effect we find is that the nuclear triplet state of the "H_2" molecule is unbound when the masses of electrons and nuclei are similar, while it is bound for large mass ratios. This strong dependence of binding on nuclear spin is absent in 3D.

Our approach can be applied to systems with dozens of particles without modifying the algorithm. More complicated sets of particles could also be treated with relatively minor changes. A very interesting direction would be to study larger systems, progressing towards 1D solids, with phonons emerging as the number of particles increase. In our approach, one would not need to make approximations in deriving an electron-phonon interaction, and one could study contributions of the phonons to entanglement entropies.

We discussed how TDVP can fail in simple situations, and present a new algorithm, a modi-

fication of TDVP, which appears to work in all situations, including long-range interactions. The key modification is the enlargement of the tangent space before each time step using global Krylov vectors. The enlarged space introduces the degrees of freedom for the correct time evolution, allowing us to combine it with the single-site TDVP method. For real-time evolution we can maintain exact unitarity, even as the bond dimension is allowed to grow. (The two-site TDVP method can be viewed as a simpler attempt to enlarge the tangent space, which may not work well with non-nearest-neighbor interactions.) The new method does not require the time step to be made particularly small, and works correctly for evolution starting with a product state. Finally, our method has excellent efficiency, with a calculation-time cost between that of one- and two-site conventional TDVP.

We expect it to be a valuable tool for out-of-equilibrium dynamics and finite-temperature simulations in systems with long-range interactions and in two dimensional systems, which has been corroborated in our study on the spin squeezing dynamics of the XXZ model with power-law decaying interactions in two dimension.

Bibliography

[1] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[2] Kenneth G. Wilson. The renormalization group: Critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47:773–840, Oct 1975.

[3] Ralf Bulla, Theo A. Costi, and Thomas Pruschke. Numerical renormalization group method for quantum impurity systems. *Rev. Mod. Phys.*, 80:395–450, Apr 2008.

[4] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.*, 90:227902, Jun 2003.

[5] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. *Phys. Rev. B*, 73:094423, Mar 2006.

[6] M B Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024–P08024, aug 2007.

[7] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82:277–306, Feb 2010.

[8] Jacob D. Bekenstein. Black holes and entropy. *Phys. Rev. D*, 7:2333–2346, Apr 1973.

[9] Mark Srednicki. Entropy and area. *Phys. Rev. Lett.*, 71:666–669, Aug 1993.

[10] Stellan Östlund and Stefan Rommer. Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.*, 75:3537–3540, Nov 1995.

[11] J Dukelsky, M. A Martín-Delgado, T Nishino, and G Sierra. Equivalence of the variational matrix product method and the density matrix renormalization group applied to spin chains. *Europhysics Letters (EPL)*, 43(4):457–462, aug 1998.

[12] R. J. Baxter, I. G. Enting, and S. K. Tsang. Hard-square lattice gas. *Journal of Statistical Physics*, 22(4):465–489, Apr 1980.

[13] Ian Affleck, Tom Kennedy, Elliott H. Lieb, and Hal Tasaki. Rigorous results on valence-bond ground states in antiferromagnets. *Phys. Rev. Lett.*, 59:799–802, Aug 1987.

[14] Ian Affleck, Tom Kennedy, Elliott H. Lieb, and Hal Tasaki. Valence bond ground states in isotropic quantum antiferromagnets. *Comm. Math. Phys.*, 115(3):477–528, 1988.

[15] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91:147902, Oct 2003.

[16] Guifré Vidal. Efficient simulation of one-dimensional quantum many-body systems. *Phys. Rev. Lett.*, 93:040502, Jul 2004.

[17] Michael P. Zaletel, Roger S. K. Mong, Christoph Karrasch, Joel E. Moore, and Frank Pollmann. Time-evolving a matrix product state with long-ranged interactions. *Phys. Rev. B*, 91:165112, Apr 2015.

[18] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pižorn, Henri Verschelde, and Frank Verstraete. Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.*, 107:070601, Aug 2011.

[19] G. Vidal. Classical simulation of infinite-size quantum lattice systems in one spatial dimension. *Phys. Rev. Lett.*, 98:070201, Feb 2007.

[20] R. Orús and G. Vidal. Infinite time-evolving block decimation algorithm beyond unitary evolution. *Phys. Rev. B*, 78:155117, Oct 2008.

[21] I. P. McCulloch. Infinite size density matrix renormalization group, revisited, 2008.

[22] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman. Variational optimization algorithms for uniform matrix product states. *Phys. Rev. B*, 97:045145, Jan 2018.

[23] F. Verstraete and J. I. Cirac. Continuous matrix product states for quantum fields. *Phys. Rev. Lett.*, 104:190405, May 2010.

[24] Frank Verstraete and J. Ignacio Cirac. Renormalization algorithms for quantum-many body systems in two and higher dimensions, 2004.

[25] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac. Classical simulation of infinite-size quantum lattice systems in two spatial dimensions. *Phys. Rev. Lett.*, 101:250602, Dec 2008.

[26] Philippe Corboz, Román Orús, Bela Bauer, and Guifré Vidal. Simulation of strongly correlated fermions in two spatial dimensions with fermionic projected entangled-pair states. *Phys. Rev. B*, 81:165104, Apr 2010.

[27] G. Vidal. Entanglement renormalization. *Phys. Rev. Lett.*, 99:220405, Nov 2007.

[28] Antoine Tilloy and J. Ignacio Cirac. Continuous tensor network states for quantum fields. *Phys. Rev. X*, 9:021040, May 2019.

[29] Jutho Haegeman, Tobias J. Osborne, Henri Verschelde, and Frank Verstraete. Entanglement renormalization for quantum fields in real space. *Phys. Rev. Lett.*, 110:100402, Mar 2013.

[30] Bo-Xiao Zheng, Chia-Min Chung, Philippe Corboz, Georg Ehlers, Ming-Pu Qin, Reinhard M. Noack, Hao Shi, Steven R. White, Shiwei Zhang, and Garnet Kin-Lic Chan. Stripe order in the underdoped region of the two-dimensional hubbard model. *Science*, 358(6367):1155–1160, 2017.

[31] Simeng Yan, David A. Huse, and Steven R. White. Spin-liquid ground state of the s = 1/2 kagome heisenberg antiferromagnet. *Science*, 332(6034):1173–1176, 2011.

[32] Michael P. Zaletel, Roger S. K. Mong, and Frank Pollmann. Topological characterization of fractional quantum hall ground states from microscopic hamiltonians. *Phys. Rev. Lett.*, 110:236801, Jun 2013.

[33] Steven R. White. Hybrid grid/basis set discretizations of the schrÃ¶dinger equation. *The Journal of Chemical Physics*, 147(24):244102, 2017.

[34] Garnet Kin-Lic Chan, Anna Keselman, Naoki Nakatani, Zhendong Li, and Steven R. White. Matrix product operators, matrix product states, and ab initio density matrix renormalization group algorithms. *The Journal of Chemical Physics*, 145(1):014102, 2016.

[35] E. Miles Stoudenmire and Steven R. White. Sliced basis density matrix renormalization group for electronic structure. *Phys. Rev. Lett.*, 119:046401, Jul 2017.

[36] Yoav Levine, Or Sharir, Nadav Cohen, and Amnon Shashua. Quantum entanglement in deep learning architectures. *Phys. Rev. Lett.*, 122:065301, Feb 2019.

[37] Edwin Stoudenmire and David J Schwab. Supervised learning with tensor networks. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29, pages 4799–4807. Curran Associates, Inc., 2016.

[38] William Huggins, Piyush Patil, Bradley Mitchell, K Birgitta Whaley, and E Miles Stoudenmire. Towards quantum machine learning with tensor networks. *Quantum Science and Technology*, 4(2):024001, jan 2019.

[39] Etienne Garand, Jia Zhou, David E. Manolopoulos, Millard H. Alexander, and Daniel M. Neumark. Nonadiabatic interactions in the cl + h2 reaction probed by clh2- and cld2- photoelectron imaging. *Science*, 319(5859):72–75, 2008.

[40] Joel M. Bowman. Beyond born-oppenheimer. *Science*, 319(5859):40–41, 2008.

[41] Graham A. Worth and Lorenz S. Cederbaum. Beyond born-oppenheimer: Molecular dynamics through a conical intersection. *Annual Review of Physical Chemistry*, 55(1):127–158, 2004. PMID: 15117250.

[42] Mingru Yang and Steven R. White. Density-matrix-renormalization-group study of a one-dimensional diatomic molecule beyond the born-oppenheimer approximation. *Phys. Rev. A*, 99:022509, Feb 2019.

[43] Adrian E. Feiguin and Steven R. White. Time-step targeting methods for real-time dynamics using the density matrix renormalization group. *Phys. Rev. B*, 72:020404(R), Jul 2005.

[44] Mingru Yang and Steven R. White. Time-dependent variational principle with ancillary krylov subspace. *Phys. Rev. B*, 102:094315, Sep 2020.

[45] Ulrich Schollwock. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96 – 192, 2011. January 2011 Special Issue.

[46] T. Barthel, M.-C. Chung, and U. Schollwöck. Entanglement scaling in critical two-dimensional fermionic and bosonic systems. *Phys. Rev. A*, 74:022329, Aug 2006.

[47] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer. Entropy, entanglement, and area: Analytical results for harmonic lattice systems. *Phys. Rev. Lett.*, 94:060503, Feb 2005.

[48] Dimitri Gioev and Israel Klich. Entanglement entropy of fermions in any dimension and the widom conjecture. *Phys. Rev. Lett.*, 96:100503, Mar 2006.

[49] Kouichi Okunishi, Yasuhiro Hieida, and Yasuhiro Akutsu. Universal asymptotic eigenvalue distribution of density matrices and corner transfer matrices in the thermodynamic limit. *Phys. Rev. E*, 59:R6227–R6230, Jun 1999.

[50] Sebastian Paeckel, Thomas Köhler, Andreas Swoboda, Salvatore R. Manmana, Ulrich Schollwöck, and Claudius Hubig. Time-evolution methods for matrix-product states. *Annals of Physics*, 411:167998, 2019.

[51] Ian P McCulloch. From density-matrix renormalization group to matrix product states. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(10):P10014–P10014, oct 2007.

[52] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix product state representations. *Quantum Info. Comput.*, 7(5):401–430, July 2007.

[53] `https://tensornetwork.org/mps/algorithms/denmat_mpo_mps/`.

[54] Bram Vanhecke, Maarten Van Damme, Jutho Haegeman, Laurens Vanderstraeten, and Frank Verstraete. Tangent-space methods for truncating uniform mps, 2020.

[55] Gregory M. Crosswhite, A. C. Doherty, and Guifré Vidal. Applying matrix product operators to model systems with long-range interactions. *Phys. Rev. B*, 78:035116, Jul 2008.

[56] Gregory M. Crosswhite and Dave Bacon. Finite automata for caching in matrix product algorithms. *Phys. Rev. A*, 78:012356, Jul 2008.

[57] B Pirvu, V Murg, J I Cirac, and F Verstraete. Matrix product operator representations. *New Journal of Physics*, 12(2):025012, feb 2010.

[58] F. Fröwis, V. Nebendahl, and W. Dür. Tensor operators: Constructions and applications for long-range interaction systems. *Phys. Rev. A*, 81:062337, Jun 2010.

[59] Steven R. White. Density matrix renormalization group algorithms with a single center site. *Phys. Rev. B*, 72:180403, Nov 2005.

[60] C. Hubig, I. P. McCulloch, U. Schollwöck, and F. A. Wolf. Strictly single-site dmrg algorithm with subspace expansion. *Phys. Rev. B*, 91:155115, Apr 2015.

[61] Steven R. White and Adrian E. Feiguin. Real-time evolution using the density matrix renormalization group. *Phys. Rev. Lett.*, 93:076401, Aug 2004.

[62] Jutho Haegeman, Christian Lubich, Ivan Oseledets, Bart Vandereycken, and Frank Verstraete. Unifying time evolution and optimization with matrix product states. *Phys. Rev. B*, 94:165116, Oct 2016.

[63] Jutho Haegeman, Tobias J. Osborne, and Frank Verstraete. Post-matrix product state methods: To tangent space and beyond. *Phys. Rev. B*, 88:075133, Aug 2013.

[64] C. Hubig, J. Haegeman, and U. Schollwöck. Error estimates for extrapolations with matrix-product states. *Phys. Rev. B*, 97:045125, Jan 2018.

[65] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. Matrix product density operators: Simulation of finite-temperature and dissipative systems. *Phys. Rev. Lett.*, 93:207204, Nov 2004.

[66] Johannes Hauschild, Eyal Leviatan, Jens H. Bardarson, Ehud Altman, Michael P. Zaletel, and Frank Pollmann. Finding purifications with minimal entanglement. *Phys. Rev. B*, 98:235163, Dec 2018.

[67] C. Karrasch, J. H. Bardarson, and J. E. Moore. Finite-temperature dynamical density matrix renormalization group and the drude weight of spin-1/2 chains. *Phys. Rev. Lett.*, 108:227206, May 2012.

[68] C Karrasch, J H Bardarson, and J E Moore. Reducing the numerical effort of finite-temperature density matrix renormalization group calculations. *New Journal of Physics*, 15(8):083031, aug 2013.

[69] Thomas Barthel. Precise evaluation of thermal response functions by optimized density matrix renormalization group schemes. *New Journal of Physics*, 15(7):073010, jul 2013.

[70] Steven R. White. Minimally entangled typical quantum states at finite temperature. *Phys. Rev. Lett.*, 102:190601, May 2009.

[71] E M Stoudenmire and Steven R White. Minimally entangled typical thermal state algorithms. *New Journal of Physics*, 12(5):055026, may 2010.

[72] Born M. and Oppenheimer R. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[73] Wheeler John Archibald. Polyelectrons. *Annals of the New York Academy of Sciences*, 48(3):219–238, 1946.

[74] Egil A. Hylleraas and Aadne Ore. Binding energy of the positronium molecule. *Phys. Rev.*, 71:493–496, Apr 1947.

[75] D. B. Cassidy and A. P. Mills Jr. The production of molecular positronium. *Nature*, 449:195 EP –, Sep 2007.

[76] Yumeng You, Xiao-Xiao Zhang, Timothy C. Berkelbach, Mark S. Hybertsen, David R. Reichman, and Tony F. Heinz. Observation of biexcitons in monolayer wse2. *Nature Physics*, 11:477 EP –, May 2015.

[77] Arne Scherrer, Federica Agostini, Daniel Sebastiani, E. K. U. Gross, and Rodolphe Vuilleumier. On the mass of atoms in molecules: Beyond the born-oppenheimer approximation. *Phys. Rev. X*, 7:031035, Aug 2017.

[78] K. Varga, J. Usukura, and Y. Suzuki. Second bound state of the positronium molecule and biexcitons. *Phys. Rev. Lett.*, 80:1876–1879, Mar 1998.

[79] J. Usukura, K. Varga, and Y. Suzuki. Signature of the existence of the positronium molecule. *Phys. Rev. A*, 58:1918–1931, Sep 1998.

[80] Y. Suzuki and J. Usukura. Excited states of the positronium molecule. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 171(1):67 – 80, 2000. Low Energy Positron and Positronium Physics.

[81] Dario Bressanini, Massimo Mella, and Gabriele Morosi. Stability of four-unit-charge systems: A quantum monte carlo study. *Phys. Rev. A*, 55:200–205, Jan 1997.

[82] Ilkka Kylänpää and Tapio T. Rantala. Thermal dissociation of dipositronium: Path-integral monte carlo approach. *Phys. Rev. A*, 80:024504, Aug 2009.

[83] Ali Abedi, Neepa T. Maitra, and E. K. U. Gross. Exact factorization of the time-dependent electron-nuclear wave function. *Phys. Rev. Lett.*, 105:123002, Sep 2010.

[84] Nikitas I. Gidopoulos and E. K. U. Gross. Electronic non-adiabatic states: towards a density functional theory beyond the born–oppenheimer approximation. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 372(2011), 2014.

[85] W. Heitler and F. London. Wechselwirkung neutraler atome und homöopolare bindung nach der quantenmechanik. *Zeitschrift für Physik*, 44(6):455–472, Jun 1927.

[86] D. M. Schrader. Symmetry of dipositronium ps_2. *Phys. Rev. Lett.*, 92:043401, Jan 2004.

[87] Donald B. Kinghorn and R. D. Poshusta. Nonadiabatic variational calculations on dipositronium using explicitly correlated gaussian basis functions. *Phys. Rev. A*, 47:3671–3681, May 1993.

[88] Matthew P. A. Fisher and Leo Radzihovsky. Quantum indistinguishability in chemical reactions. *Proceedings of the National Academy of Sciences*, 115(20):E4551–E4558, 2018.

[89] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, Oct 1993.

[90] Werner Kutzelnigg. The adiabatic approximation i. the physical background of the born-handy ansatz. *Molecular Physics*, 90(6):909–916, 1997.

[91] Thomas E. Baker, E. Miles Stoudenmire, Lucas O. Wagner, Kieron Burke, and Steven R. White. One-dimensional mimicking of electronic structure: The case for exponentials. *Phys. Rev. B*, 91:235141, Jun 2015.

[92] I. P. McCulloch and M. GulÃ¡csi. The non-abelian density matrix renormalization group algorithm. *EPL (Europhysics Letters)*, 57(6):852, 2002.

[93] Kieron Burke Lucas O. Wagner, E. M. Stoudenmire and Steven R. White. Reference electronic structure calculations in one dimension. *Phys. Chem. Chem. Phys.*, 14:8581–8590, 2012.

[94] Seung Kyu Min, Ali Abedi, Kwang S. Kim, and E. K. U. Gross. Is the molecular berry phase an artifact of the born-oppenheimer approximation? *Phys. Rev. Lett.*, 113:263004, Dec 2014.

[95] E.M. Stoudenmire and Steven R. White. Studying two-dimensional systems with the density matrix renormalization group. *Annual Review of Condensed Matter Physics*, 3(1):111–128, 2012.

[96] P. E. Dargel, A. Wöllert, A. Honecker, I. P. McCulloch, U. Schollwöck, and T. Pruschke. Lanczos algorithm with matrix product states for dynamical correlation functions. *Phys. Rev. B*, 85:205119, May 2012.

[97] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch. Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional bose gas. *Nature Physics*, 8(4):325–330, Apr 2012.

[98] Enrico Ronca, Zhendong Li, Carlos A. Jimenez-Hoyos, and Garnet Kin-Lic Chan. Time-step targeting time-dependent and dynamical density matrix renormalization group algorithms with ab initio hamiltonians. *Journal of Chemical Theory and Computation*, 13(11):5560–5571, 2017.

[99] P. A. M. Dirac. Note on exchange phenomena in the thomas atom. *Mathematical Proceedings of the Cambridge Philosophical Society*, 26(3):376–385, 1930.

[100] Kévin Hémery, Frank Pollmann, and David J. Luitz. Matrix product states approaches to operator spreading in ergodic quantum systems. *Phys. Rev. B*, 100:104303, Sep 2019.

[101] Benedikt Kloss, David R. Reichman, and Yevgeny Bar Lev. Studying dynamics in two-dimensional quantum lattices using tree tensor network states. *SciPost Phys.*, 9:70, 2020.

[102] Shimpei Goto and Ippei Danshita. Performance of the time-dependent variational principle for matrix product states in the long-time evolution of a pure state. *Phys. Rev. B*, 99:054307, Feb 2019.

[103] Sergey V. Dolgov and Dmitry V. Savostyanov. Alternating minimal energy methods for linear systems in higher dimensions. *SIAM Journal on Scientific Computing*, 36(5):A2248–A2271, 2014.

[104] Y. Saad. Analysis of some krylov subspace approximations to the matrix exponential operator. *SIAM Journal on Numerical Analysis*, 29(1):209–228, 1992.

[105] Roger B. Sidje. Expokit: A software package for computing matrix exponentials. *ACM Trans. Math. Softw.*, 24(1):130–156, March 1998.

[106] M L Wall and Lincoln D Carr. Out-of-equilibrium dynamics with matrix product states. *New Journal of Physics*, 14(12):125015, dec 2012.

[107] Masahiro Kitagawa and Masahito Ueda. Squeezed spin states. *Phys. Rev. A*, 47:5138–5143, Jun 1993.

[108] Jian Ma, Xiaoguang Wang, C. P. Sun, and Franco Nori. Quantum spin squeezing. *Physics Reports*, 509(2):89 – 165, 2011.

[109] Michael A. Perlin, Chunlei Qu, and Ana Maria Rey. Spin squeezing with short-range spin-exchange interactions, 2020.

[110] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A*, 46:R6797–R6800, Dec 1992.

[111] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen. Squeezed atomic states and projection noise in spectroscopy. *Phys. Rev. A*, 50:67–88, Jul 1994.

[112] D.F. Walls and P. Zoller. Enhanced sensitivity of a gravitational wave detector. *Physics Letters A*, 85(2):118 – 120, 1981.

[113] J. Aasi et al. Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nature Photonics*, 7(8):613–619, Aug 2013.

[114] M. Kitagawa and Y. Yamamoto. Number-phase minimum-uncertainty state with reduced number uncertainty in a kerr nonlinear interferometer. *Phys. Rev. A*, 34:3974–3988, Nov 1986.

[115] Y. Takahashi, K. Honda, N. Tanaka, K. Toyoda, K. Ishikawa, and T. Yabuzaki. Quantum nondemolition measurement of spin via the paramagnetic faraday rotation. *Phys. Rev. A*, 60:4974–4979, Dec 1999.

[116] Michael Foss-Feig, Zhe-Xuan Gong, Alexey V. Gorshkov, and Charles W. Clark. Entanglement and spin-squeezing without infinite-range interactions, 2016.

[117] A. M. Rey, L. Jiang, M. Fleischhauer, E. Demler, and M. D. Lukin. Many-body protected entanglement generation in interacting spin systems. *Phys. Rev. A*, 77:052305, May 2008.

[118] P. Cappellaro and M. D. Lukin. Quantum correlation in disordered spin systems: Applications to magnetic sensing. *Phys. Rev. A*, 80:032311, Sep 2009.

[119] P. He, M. A. Perlin, S. R. Muleady, R. J. Lewis-Swan, R. B. Hutson, J. Ye, and A. M. Rey. Engineering spin squeezing in a 3d optical lattice with interacting spin-orbit-coupled fermions. *Phys. Rev. Research*, 1:033075, Nov 2019.

[120] J. Schachenmayer, A. Pikovski, and A. M. Rey. Many-body quantum spin dynamics with monte carlo trajectories on a discrete phase space. *Phys. Rev. X*, 5:011022, Feb 2015.

[121] Bihui Zhu, Ana Maria Rey, and Johannes Schachenmayer. A generalized phase space approach for solving quantum spin dynamics. *New Journal of Physics*, 21(8):082001, aug 2019.

[122] Anatoli Polkovnikov. Phase space representation of quantum dynamics. *Annals of Physics*, 325(8):1790 – 1852, 2010.

[123] P. Bruno. Absence of spontaneous magnetic order at nonzero temperature in one- and two-dimensional heisenberg and XY systems with long-range interactions. *Phys. Rev. Lett.*, 87:137203, Sep 2001.

[124] M. P. Kwasigroch and N. R. Cooper. Bose-einstein condensation and many-body localization of rotational excitations of polar molecules following a microwave pulse. *Phys. Rev. A*, 90:021605, Aug 2014.

[125] Irénée Frérot, Piero Naldesi, and Tommaso Roscilde. Entanglement and fluctuations in the xxz model with power-law interactions. *Phys. Rev. B*, 95:245111, Jun 2017.

[126] Lucas O. Wagner, E. M. Stoudenmire, Kieron Burke, and Steven R. White. Guaranteed convergence of the kohn-sham equations. *Phys. Rev. Lett.*, 111:093003, Aug 2013.

[127] Michele Dolfi, Bela Bauer, Matthias Troyer, and Zoran Ristivojevic. Multigrid algorithms for tensor network states. *Phys. Rev. Lett.*, 109:020604, Jul 2012.

[128] Glen Evenbly and Steven R. White. Entanglement renormalization and wavelets. *Phys. Rev. Lett.*, 116:140403, Apr 2016.

[129] Alfred Haar. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3):331–371, Sep 1910.

[130] Glen Evenbly and Steven R. White. Representation and design of wavelets using unitary circuits. *Phys. Rev. A*, 97:052314, May 2018.

[131] Natalia Chepiga and Steven R. White. Comb tensor networks. *Phys. Rev. B*, 99:235426, Jun 2019.

Appendix A

Details of the calculations in the problem of a one-dimensional diatomic molecule beyond the Born-Oppenheimer approximation

A.1 Generalization of the Hartree-Fock approximation to go beyond the Born-Oppenheimer approximation

The Hamiltonian for a one-dimensional system of N nuclei with charge $+e$ and N electrons with charge $-e$ reads

$$\begin{aligned} H = & \int \mathrm{d}x \sum_{\alpha,s} \phi^{\dagger}_{\alpha,s}(x) \left[-\frac{\hbar^2}{2m_\alpha} \frac{\mathrm{d}^2}{\mathrm{d}x^2} \right] \phi_{\alpha,s}(x) \\ & + \frac{1}{2} \int \mathrm{d}x \mathrm{d}y \sum_{\alpha\beta} V_{\alpha\beta}(x-y) \sum_{ss'} \phi^{\dagger}_{\alpha,s}(x) \phi^{\dagger}_{\beta,s'}(y) \phi_{\beta,s'}(y) \phi_{\alpha,s}(x), \end{aligned} \tag{A.1}$$

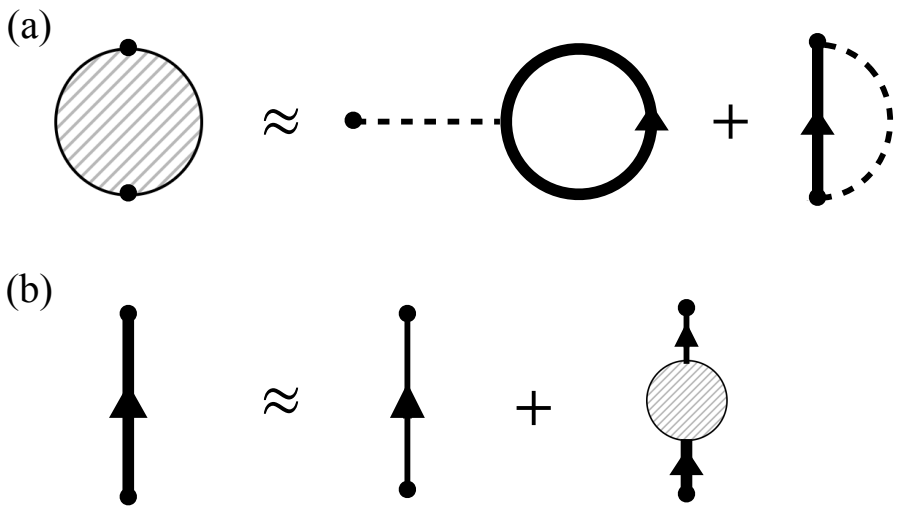

Figure A.1: (A) Diagram for the self-consistent self energy in the Hartree-Fock approximation. (b) Diagram for the self-consistent Green's function in the Hartree-Fock approximation.

where $\alpha \in \{p, e\}$ denotes the nuclei or electrons and s the spin index, and $V_{\alpha\beta}(x-y) = V(x-y)M_{\alpha\beta}$ and M takes the form

$$M = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \tag{A.2}$$

The Hartree-Fock approximation assumes that each particle moves in a single-particle potential that comes from its average interaction with all of the other particles. Then the single-particle energy should be the unperturbed energy plus the potential energy of interaction averaged over the states occupied by all of the other particles. In field theory language, the Hartree-Fock approximation corresponds to solve the Dyson-Schwinger equation for the Green's function and the proper self energy self-consistently, with diagrams to the first order, i.e. including only the direct interaction and the exchange interaction.

To go beyond the Born-Oppenheimer approximation, to treat the two species of fermions, we only need to treat the species as a degree of freedom similar to spin and allow the interaction to be species dependent. From now on, we use x to denote $(t, \mathbf{x})$. The Hartree-

Fock approximation (FIG. A.1) for this system gives the self-consistent Green's function

$$G_{\alpha\beta,s\sigma}(x,y) \approx G^0_{\alpha\beta,s\sigma}(x,y) + \int \mathrm{d}x_1 \mathrm{d}x_1' \sum_{\lambda\gamma,\eta\nu} G^0_{\alpha\lambda,s\eta}(x,x_1)\Sigma^*_{\lambda\gamma,\eta\nu}(x_1,x_1')G_{\gamma\beta,\nu\sigma}(x_1',y), \quad \text{(A.3)}$$

and the self-consistent proper self energy

$$\begin{aligned}\Sigma^*_{\lambda\gamma,\eta\nu}(x_1,x_1') \approx &-\frac{\mathrm{i}}{\hbar}\delta(x_1-x_1')\int \mathrm{d}x_2 \sum_{\lambda'\mu\mu',\eta'\rho\rho'} U_{\lambda\eta\lambda'\eta',\mu\rho\mu'\rho'}(x_1,x_2)G_{\mu\mu',\rho\rho'}(x_2,x_2)\delta_{\lambda'\gamma}\delta_{\eta'\nu} \\ &+\frac{\mathrm{i}}{\hbar}\sum_{\lambda'\mu\mu',\eta'\rho\rho'} U_{\lambda\eta\lambda'\eta',\mu\rho\mu'\rho'}(x_1,x_1')G_{\lambda'\mu,\eta'\rho}(x_1,x_1')\delta_{\mu'\gamma}\delta_{\rho'\nu}\end{aligned} \quad \text{(A.4)}$$

where $U_{\lambda\eta\lambda'\eta',\mu\rho\mu'\rho'}(x_1,x_1') = \delta_{\eta\eta'}\delta_{\rho\rho'}V(x_1-x_1')M_{\lambda\mu}\delta_{\lambda\lambda'}\delta_{\mu\mu'}$ and the Green's function should be diagonal in both the spin index and the species index, due to charge conservation and spin conservation in each species.

In the restricted Hartree-Fock, we requires the same orbital for different spins but allows different orbitals for different species. So we have

$$G_{\alpha\beta,s\sigma}(x,y) = G_\alpha(x,y)\delta_{\alpha\beta}\delta_{s\sigma}, \quad \text{(A.5)}$$

and similarly for $G^0_{\alpha\beta,s\sigma}(x,y)$. The expressions of G and Σ^* are simplified to

$$G_\alpha(x,y) \approx G^0_\alpha(x,y) + \int \mathrm{d}x_1 \mathrm{d}x_1' G^0_\alpha(x,x_1)\Sigma^*_\alpha(x_1,x_1')G_\alpha(x_1',y) \quad \text{(A.6)}$$

and

$$\begin{aligned}\Sigma^*_\alpha(x_1,x_1') \approx -\frac{\mathrm{i}}{\hbar}\delta(t_1-t_1')\Big[&(2s+1)\delta(\mathbf{x}_1-\mathbf{x}_1')\int \mathrm{d}\mathbf{x}_2 \sum_\mu M_{\alpha\mu}G_\mu(x_2,x_2)V(\mathbf{x}_1-\mathbf{x}_2) \\ &- V(\mathbf{x}_1-\mathbf{x}_1')G_\alpha(x_1,x_1')\Big],\end{aligned} \quad \text{(A.7)}$$

where $s = 1/2$. Fourier transform in time gives

$$G_\alpha(\mathbf{x},\mathbf{y};\omega) \approx G^0_\alpha(\mathbf{x},\mathbf{y};\omega) + \int \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}'_1 G^0_\alpha(\mathbf{x},\mathbf{x}_1;\omega)\Sigma^*_\alpha(\mathbf{x}_1,\mathbf{x}'_1)G_\alpha(\mathbf{x}'_1,\mathbf{y};\omega) \tag{A.8}$$

and

$$\Sigma^*_\alpha(\mathbf{x}_1,\mathbf{x}'_1) \approx -\frac{\mathrm{i}}{\hbar}\Big[(2s+1)\delta(\mathbf{x}_1-\mathbf{x}'_1)\int \mathrm{d}\mathbf{x}_2 V(\mathbf{x}_1-\mathbf{x}_2)\sum_\mu M_{\alpha\mu}\int \frac{\mathrm{d}\omega}{2\pi}e^{\mathrm{i}\omega\eta}G_\mu(\mathbf{x}_2,\mathbf{x}_2;\omega) - V(\mathbf{x}_1-\mathbf{x}'_1)\int \frac{\mathrm{d}\omega}{2\pi}e^{\mathrm{i}\omega\eta}G_\alpha(\mathbf{x}_1,\mathbf{x}'_1;\omega)\Big]. \tag{A.9}$$

Consider a complete set of single-particle orthonormal eigenstates of H_0, which is spin-independent,

$$\left[-\frac{\hbar^2\nabla^2}{2m_\alpha}\right]\phi^0_{j,\alpha}(\mathbf{x}) = \epsilon^0_{j,\alpha}\phi^0_{j,\alpha}(\mathbf{x}). \tag{A.10}$$

The non-interacting Green's function becomes

$$\mathrm{i}G^0_\alpha(x,x') = \sum_j \phi^0_{j,\alpha}(\mathbf{x})\phi^{0*}_{j,\alpha}(\mathbf{x}')e^{-\mathrm{i}\epsilon^0_{j,\alpha}(t-t')/\hbar}\Big[\theta(t-t')\theta(\epsilon^0_{j,\alpha}-\epsilon^0_{F,\alpha})-\theta(t'-t)\theta(\epsilon^0_{F,\alpha}-\epsilon^0_{j,\alpha})\Big], \tag{A.11}$$

where $\epsilon^0_{F,\alpha}$ is the energy of the last filled single-particle state for α. The Fourier transform gives

$$G^0_\alpha(\mathbf{x},\mathbf{x}';\omega) = \sum_j \phi^0_{j,\alpha}(\mathbf{x})\phi^{0*}_{j,\alpha}(\mathbf{x}')\left[\frac{\theta(\epsilon^0_{j,\alpha}-\epsilon^0_{F,\alpha})}{\omega-\epsilon^0_{j,\alpha}/\hbar+\mathrm{i}\eta}+\frac{\theta(\epsilon^0_{F,\alpha}-\epsilon^0_{j,\alpha})}{\omega-\epsilon^0_{j,\alpha}/\hbar-\mathrm{i}\eta}\right]. \tag{A.12}$$

Hartree-Fock seeks a solution for G_α in the same form as G^0_α, i.e.

$$G_\alpha(\mathbf{x},\mathbf{x}';\omega) = \sum_j \phi_{j,\alpha}(\mathbf{x})\phi^*_{j,\alpha}(\mathbf{x}')\left[\frac{\theta(\epsilon_{j,\alpha}-\epsilon_{F,\alpha})}{\omega-\epsilon_{j,\alpha}/\hbar+\mathrm{i}\eta}+\frac{\theta(\epsilon_{F,\alpha}-\epsilon_{j,\alpha})}{\omega-\epsilon_{j,\alpha}/\hbar-\mathrm{i}\eta}\right], \tag{A.13}$$

where $\{\phi_{j,\alpha}(\mathbf{x})\}$ denotes a complete set of single-particle wavefunctions with energy $\epsilon_{j,\alpha}$. The Dyson equation for the Green's function then becomes the self-consistent equation for the single-particle wavefunctions in the restricted Hartree-Fock

$$\left[-\frac{\hbar^2\nabla^2}{2m_\alpha}\right]\phi_{j,\alpha}(\mathbf{x}) + \int \mathrm{d}\mathbf{x}'\hbar\Sigma^*_\alpha(\mathbf{x},\mathbf{x}')\phi_{j,\alpha}(\mathbf{x}') = \epsilon_{j\alpha}\phi_{j,\alpha}(\mathbf{x}), \tag{A.14}$$

where

$$\begin{aligned}\hbar\Sigma^*_\alpha(\mathbf{x},\mathbf{x}') =& \delta(\mathbf{x}-\mathbf{x}')\int \mathrm{d}\mathbf{x}_1 V(\mathbf{x}-\mathbf{x}_1)\sum_\mu M_{\alpha\mu}n_\mu(\mathbf{x}_1) \\ &- V(\mathbf{x}-\mathbf{x}')\sum_j \phi_{j,\alpha}(\mathbf{x})\phi^*_{j,\alpha}(\mathbf{x}')\theta(\epsilon_{F,\alpha}-\epsilon_{j,\alpha}),\end{aligned} \tag{A.15}$$

which consists of a local direct term proportional to the particle density $n_\mu(\mathbf{x}) = (2s+1)\sum_j |\phi_{j,\mu}(\mathbf{x})|^2\theta(\epsilon_{F,\mu}-\epsilon_{j,\mu})$ and a nonlocal exchange term.

In the unrestricted Hartree-Fock, we allows different orbitals both for different spins and different species. So we have

$$G_{\alpha\beta,s\sigma}(x,y) = G_{\alpha,s}(x,y)\delta_{\alpha\beta}\delta_{s\sigma}, \tag{A.16}$$

and the single particle states are also both spin- and species-dependent. Similarly, we have

$$\left[-\frac{\hbar^2\nabla^2}{2m_\alpha}\right]\phi_{j,\alpha,s}(\mathbf{x}) + \int \mathrm{d}\mathbf{x}'\hbar\Sigma^*_{\alpha,s}(\mathbf{x},\mathbf{x}')\phi_{j,\alpha,s}(\mathbf{x}') = \epsilon_{j\alpha s}\phi_{j,\alpha,s}(\mathbf{x}) \tag{A.17}$$

for the unrestricted Hartree-Fock, where

$$\begin{aligned}\hbar\Sigma^*_{\alpha,s}(\mathbf{x},\mathbf{x}') =& \delta(\mathbf{x}-\mathbf{x}')\int \mathrm{d}\mathbf{x}_1 V(\mathbf{x}-\mathbf{x}_1)\sum_\mu M_{\alpha\mu}n_\mu(\mathbf{x}_1) \\ &- V(\mathbf{x}-\mathbf{x}')\sum_{j'} \phi_{j',\alpha,s}(\mathbf{x})\phi^*_{j',\alpha,s}(\mathbf{x}')\theta(\epsilon_{F,\alpha,s}-\epsilon_{j',\alpha,s})\end{aligned} \tag{A.18}$$

with $n_\mu(\mathbf{x}) = \sum_{\sigma,j} |\phi_{j,\mu,\sigma}(\mathbf{x})|^2 \theta(\epsilon_{F,\mu,\sigma} - \epsilon_{j,\mu,\sigma})$.

For the unrestricted Hartree-Fock, we define

$$\hat{h}_\alpha(\mathbf{x}) = -\frac{\hbar^2}{2m_\alpha}\frac{\mathrm{d}^2}{\mathrm{d}x^2} \tag{A.19}$$

and the Coulomb operator

$$\hat{J}_\alpha(\mathbf{x}) = \int \mathrm{d}\mathbf{x}' V(\mathbf{x}-\mathbf{x}') \sum_\mu M_{\alpha\mu} n_\mu(\mathbf{x}') \tag{A.20}$$

and the exchange operator

$$\hat{K}_{j',\alpha s}(\mathbf{x}) = \int \mathrm{d}\mathbf{x}' V(\mathbf{x}-\mathbf{x}') \phi_{j',\alpha,s}(\mathbf{x}) \phi^*_{j',\alpha,s}(\mathbf{x}') \theta(\epsilon_{F,\alpha,s} - \epsilon_{j',\alpha,s}), \tag{A.21}$$

which nontrivially acts on a not-yet-converged $\tilde{\phi}_{j,\alpha,s}(\mathbf{x})$, i.e.

$$\hat{K}_{j',\alpha s}(\mathbf{x})\tilde{\phi}_{j,\alpha,s}(\mathbf{x}) = \left[\int \mathrm{d}\mathbf{x}' V(\mathbf{x}-\mathbf{x}') \tilde{\phi}_{j,\alpha,s}(\mathbf{x}') \phi^*_{j',\alpha,s}(\mathbf{x}') \theta(\epsilon_{F,\alpha,s} - \epsilon_{j',\alpha,s})\right] \phi_{j',\alpha,s}(\mathbf{x}). \tag{A.22}$$

So the Fock operator is given by

$$\hat{F}_{\alpha s}(\mathbf{x}) = \hat{h}_\alpha(\mathbf{x}) + \hat{J}_\alpha(\mathbf{x}) - \sum_{j'} \hat{K}_{j',\alpha s}(\mathbf{x}). \tag{A.23}$$

Therefore, solving Eq. (A.17) for a given self energy $\Sigma^*_{\alpha,s}$ is equivalent to solving the eigenvalue problem

$$\hat{F}_{\alpha s}(\mathbf{x})\phi_{j,\alpha,s}(\mathbf{x}) = \epsilon_{j\alpha s}\phi_{j,\alpha,s}(\mathbf{x}). \tag{A.24}$$

The generalized unrestricted Hartree-Fock proceeds[1] as follows:

[1]For good convergence, we use the method introduced in Ref. [126]

1. Initialize $\phi_{j,\alpha,s}(\mathbf{x})$ with the non-interacting single-particle eigenfunctions $\phi^0_{j,\alpha,s}(\mathbf{x})$, for $j = 1, \ldots, F$;

2. Calculate the Fock operator from $\phi_{j,\alpha,s}(\mathbf{x})$;

3. Given the Fock operator calculated in step (2), for $j = 1, \ldots, F$, solve the eigenvalue problem in Eq. (A.24) and get a new $\tilde{\phi}_{j,\alpha,s}(\mathbf{x})$ corresponding to the single-particle eigenenergy $\tilde{\epsilon}_{j\alpha s}$. If a convergence in energy has been reached, stop the iterations, otherwise go back to step (2).

In the one-dimensional diatomic molecule problem, we have two electrons and two nuclei. If we want to get the Hartree-Fock state for the singlet-singlet state, we only need to solve $\phi_{1,p,\uparrow}(\mathbf{x})$, $\phi_{1,p,\downarrow}(\mathbf{x})$, $\phi_{1,e,\uparrow}(\mathbf{x})$, $\phi_{1,e,\downarrow}(\mathbf{x})$, since we have only one particle for each α, s. If the nuclei are in $|\uparrow\uparrow\rangle$ state and the electrons are in singlet, then we need to solve $\phi_{1,p,\uparrow}(\mathbf{x})$, $\phi_{2,p,\uparrow}(\mathbf{x})$, $\phi_{1,e,\uparrow}(\mathbf{x})$, $\phi_{1,e,\downarrow}(\mathbf{x})$. To compare with DMRG, which is of finite size, we also use a hard wall boundary condition at edges for the Hartree-Fock calculation. So the non-interacting single particle eigenfunctions can be taken as the eigenfunctions for the problem of one particle in a one-dimensional infinitely deep square well.

We still need to calculate the ground state energy

$$E = -\frac{1}{2}\mathrm{i}\int \mathrm{d}\mathbf{x} \lim_{t'\to t^+} \lim_{\mathbf{x}'\to\mathbf{x}} \sum_{\alpha,s} \left[\mathrm{i}\hbar\frac{\partial}{\partial t} - \frac{\hbar^2\nabla^2}{2m_\alpha}\right] G_{\alpha,s}(x, x') \tag{A.25}$$

or in terms of the single-particle wavefunctions

$$E = \sum_{j,\alpha,s} \theta(\epsilon_{F,\alpha,s} - \epsilon_{j,\alpha,s}) \left[\epsilon_{j,\alpha,s} - \frac{1}{2}\int \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' \phi^*_{j,\alpha,s}(\mathbf{x})\hbar\Sigma^*_{\alpha,s}(\mathbf{x}, \mathbf{x}')\phi_{j,\alpha,s}(\mathbf{x}')\right] \tag{A.26}$$

To work with lattice, we need to discretize the unrestricted Hartree-Fock equation. The

fourth order finite difference formula gives

$$\frac{\mathrm{d}^2\phi_{j,\alpha,s}(\mathbf{x})}{\mathrm{d}x^2} \approx \frac{1}{12(\Delta x)^2}\sum_{q=1}^{N_L}\big[-\delta_{p+2,q}+16\delta_{p+1,q}-30\delta_{p,q}+16\delta_{p-1,q}-\delta_{p-2,q}\big]\phi_{j,\alpha,s;q}, \quad \text{(A.27)}$$

where N_L is the lattice size. We can also discretize the integral in the Coulomb and exchange operator, and the unrestricted Hartree-Fock equation becomes

$$\sum_q (h_{\alpha;pq}+J_{\alpha;p}\delta_{pq}-K_{\alpha,s;pq})\phi_{j,\alpha,s;q}=\epsilon_{j\alpha s}\phi_{j,\alpha,s;p}, \quad \text{(A.28)}$$

which is a matrix equation with

$$\begin{aligned} h_{\alpha;pq} &= -\frac{\hbar^2}{2m_\alpha}\frac{1}{12(\Delta x)^2}\big[-\delta_{p+2,q}+16\delta_{p+1,q}-30\delta_{p,q}+16\delta_{p-1,q}-\delta_{p-2,q}\big], \\ J_{\alpha;p} &= \Delta x\sum_{q'}V_{pq'}\sum_{\mu}M_{\alpha\mu}n_{\mu;q'}, \\ K_{\alpha,s;pq} &= \Delta x V_{pq}n_{\alpha,s;pq}, \end{aligned} \quad \text{(A.29)}$$

where $V_{pq}=V(\mathbf{x}_p-\mathbf{x}_q)$, $n_{\mu;q'}=n_\mu(\mathbf{x}_{q'})$, and $n_{\alpha,s;pq}=\sum_{j'}\phi_{j',\alpha,s}(\mathbf{x}_p)\phi^*_{j',\alpha,s}(\mathbf{x}_q)\theta(\epsilon_{F,\alpha,s}-\epsilon_{j',\alpha,s})$. The ground state energy in Eq. (A.26) can also be discretized similarly.

A.2 Tricks to facilitate the convergence of DMRG for a system of two energy scales

When the mass ratio M between the two species of particles are very large, the kinetic energy becomes much smaller for one type of particle than the other, which makes it difficult for DMRG to converge especially when the number of particles $\gtrsim 10$. The simplest thing one could do is to start with a better initial state. For example the ground state obtained in the BO limit will be a good ansatz. One can also start with small mass ratios and slowly

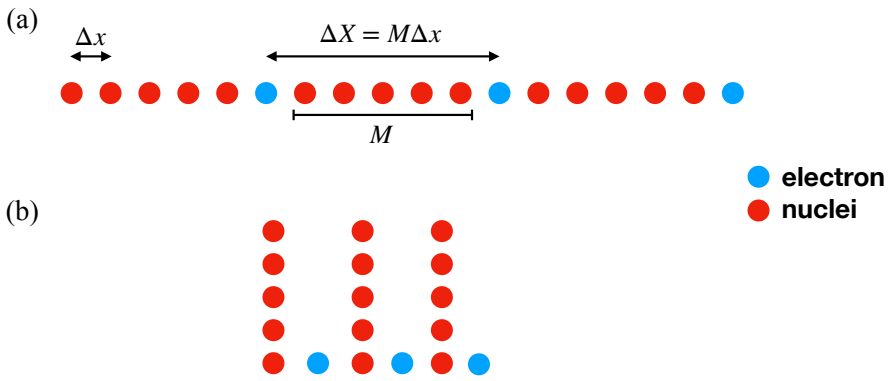

Figure A.2: (a) Geometry of the MPS if we make the electron's lattice spacing M times large than the nuclei. (b) Geometry of the comb tensor network.

increase the mass ratio in the DMRG sweeps.

One technique to facilitate the convergence of DMRG for systems with multiple length scales is called the multigrid algorithm[127]. Suppose we would like to get results on a grid with a lattice spacing Δx. In the algorithm, we start with a coarser grid and do DMRG. Then we interpolate the wavefunction to a finer grid and do DMRG again. We repeat this procedure until we reach the targeted fine grid with the lattice spacing Δx. This interpolation can be done by applying an isometry[27] $\hat{T}^{\tilde{\sigma}}_{\sigma_1 \ldots \sigma_n}$ to each site tensor of the MPS, followed by a SVD, if we want to expand 1 site to n sites. This isometry can also be understood as a wavelet transformation[128]. The simplest isometry is then a Haar gate[129]. Higher order wavelet transformations can be applied to improve the smoothness of the interpolation[130]. For fermionic systems, the signs due to permutation of fermions need to be taken care of for the wavelet transformations. However, we found that this method does not work very well. Our experience is that after interpolating to a finer grid, we cannot do many sweeps until the DMRG gets stuck in a local minimum.

Alternatively, it would be natural to use different lattice spacing for nuclei and electrons when discretizing the system, since the lattice spacing should be irrelevant to the low-energy physics as long as it is fine enough. Then the number of nuclei sites will be M times of the

electron sites. Since there is no hopping between electrons and nuclei, we actually face a worse situation in getting trapped in a local minimum. To solve the problem, one could use swap gates to permute the electron sites to be adjacent and do a local DMRG update and then permute it back (FIG. A.2(a)). Alternately, we can also use a tree tensor network which has the comb[131] geometry (FIG. A.2(b)), so the number of swap operations needed will be much smaller. We found this method works much better than the multigrid algorithm in improving the convergence of DMRG.

Appendix B

Details of the GSE-TDVP algorithm

B.1 Proof of the equivalence between the MPS and the density matrix formulation

In the following we prove that SVD C_i' in Eq. (3.2) is equivalent to diagonalizing the sum of the right reduced density matrix $\rho_i + \tilde{\rho}_i$.

At site N, if $C'_N = U'_N S'_N V'_N$, then

$$\begin{aligned} \rho'_N = {C'_N}^\dagger C'_N &= \begin{bmatrix} C_N^\dagger & \tilde{C}_N^\dagger \end{bmatrix} \begin{bmatrix} C_N \\ \tilde{C}_N \end{bmatrix} \\ &= C_N^\dagger C_N + \tilde{C}_N^\dagger \tilde{C}_N = \rho_N + \tilde{\rho}_N \\ &= {V'_N}^\dagger {S'_N}^2 V'_N. \end{aligned} \tag{B.1}$$

So at site N, we get the same V'_N in diagonalization of $\rho_N + \tilde{\rho}_N$ as in SVD of C'_N. We can

write U'_N into a block form

$$U'_N = \begin{bmatrix} U_N \\ \tilde{U}_N \end{bmatrix} \tag{B.2}$$

so $C_N = U_N S'_N V'_N$ and $\tilde{C}_N = \tilde{U}_N S'_N V'_N$. Absorbing $U'_N S'_N$ into A'_{N-1} is equivalent to moving the orthogonality center to site $N-1$ separately for $|\psi\rangle$ and $|\tilde{\psi}\rangle$, i.e.

$$\begin{aligned} C'_{N-1} &= A'_{N-1} U'_N S'_N \\ &= \begin{bmatrix} A_{N-1} & 0 \\ 0 & \tilde{A}_{N-1} \end{bmatrix} \begin{bmatrix} U_N \\ \tilde{U}_N \end{bmatrix} S'_N \\ &= \begin{bmatrix} A_{N-1} U_N S'_N \\ \tilde{A}_{N-1} \tilde{U}_N S'_N \end{bmatrix} = \begin{bmatrix} A_{N-1} C_N {V'_N}^\dagger \\ \tilde{A}_{N-1} \tilde{C}_N {V'_N}^\dagger \end{bmatrix} \\ &= \begin{bmatrix} C_{N-1} \\ \tilde{C}_{N-1} \end{bmatrix}. \end{aligned} \tag{B.3}$$

Similar to B.1, SVD C'_{N-1} is equivalent to diagonalizing $\rho_{N-1} + \tilde{\rho}_{N-1}$ and so on for all other sites. This proof can be extended to the cases of $k > 2$ MPS's.

B.2 Global subspace expansion

In the following we summarize the step 2 of the global subspace expansion. It deals with how to extend the basis of the left canonical MPS of $|\psi\rangle$ by the left canonical MPS's of $|\tilde{\psi}_1\rangle, \ldots, |\tilde{\psi}_{k-1}\rangle$.

The ith iteration is described as follows:

1. Form the one-site right reduced density matrix ρ_i, $\tilde{\rho}_{1,i}$, $\ldots$, $\tilde{\rho}_{k-1,i}$ from the orthogonality

center C_i, $\tilde{C}_{1,i}$, ..., $\tilde{C}_{k-1,i}$ at site i for $|\psi\rangle$, $|\tilde{\psi}_1\rangle$, ..., $|\tilde{\psi}_{k-1}\rangle$ respectively.

2. Diagonalize ρ_i and get $\rho_i = V_i^\dagger S_i^2 V_i$. Not truncate here.

3. Form a projection operator onto the null space of C_i, i.e. $P_i = \mathbb{1} - V_i^\dagger V_i$.

4. If $P_i \neq 0$, do the summation $\tilde{\rho}_i = \tilde{\rho}_{1,i} + \cdots + \tilde{\rho}_{k-1,i}$, and project $\tilde{\rho}_i$ by $\tilde{\rho}_i^\perp = P_i \tilde{\rho}_i P_i$.

5. Diagonalize $\tilde{\rho}_i^\perp$ and get $\tilde{\rho}_i^\perp = \tilde{V}_i^{\perp\dagger} \tilde{S}_i^{\perp 2} \tilde{V}_i^\perp$. Truncate η weights.

6. Enlarge the row space of V_i by direct sum with $\tilde{V}_i^\perp$ and get $V_i' = [V_i \ \tilde{V}_i^\perp]^T$.

7. Multiply $V_i'^\dagger$ with $A_{i-1}C_i$, $\tilde{A}_{1,i-1}\tilde{C}_{1,i}$, ..., $\tilde{A}_{k-1,i-1}\tilde{C}_{k-1,i}$ respectively and get the next orthogonality center C_i, $\tilde{C}_{1,i-1}$, ..., $\tilde{C}_{k-1,i-1}$ at site $i-1$.

B.3 Local subspace expansion

Local subspace expansion has the advantage that edge tensors are not reconstructed each time after a subspace expansion. However, the accuracy improved by this way is worse than the global subspace expansion given the same enlarged bond dimension and higher order of application of $L_{i-1}W_i$ does not further improve the accuracy.

In the following we will take the left-to-right single-site TDVP sweep as an example to illustrate the local subspace expansion and the adjoint right-to-left sweep can be formulated similarly.

Just like the subspace expansion for the single-site DMRG, we use the expansion term with dimension $(m_{i-1}, d_i, w_i m_i)$

$$R_i = \alpha L_{i-1} W_i C_i. \tag{B.4}$$

where α is a constant to be tuned, L_{i-1} is the left edge tensor of the local effective Hamiltonian $H(i)$, and W_i is the tensor of the MPO of $\hat{H}$ at site i.

The ith iteration is as follows:

1. Integrate the equation $\mathrm{i}\frac{\mathrm{d}C_i(t)}{\mathrm{d}t} = H(i)C_i(t)$ and get $C_i(t+\Delta t)$.
2. SVD C_i and get $C_i = U_i S_i V_i$. Not truncate here.
3. Form a projector to the orthogonal complement of the column space of C_i, i.e. $P_i = \mathbb{1} - U_i U_i^\dagger$.
4. If $P_i \neq 0$, project $R_i = \alpha L_{i-1} W_i C_i$ by $R_i^\perp = P_i R_i$.
5. SVD $R_i^\perp$ and get $R_i^\perp = U_i^\perp S_i^\perp V_i^\perp$. Truncate if necessary.
6. Enlarge the column space of U_i by direct sum with $U_i^\perp$, i.e. $U'_i = [U_i \;\; U_i^\perp]$.
7. Multiply $U'^\dagger_i$ with C_i and get $D_i = U'^\dagger_i C_i$.
8. Integrate the equation $-\mathrm{i}\frac{\mathrm{d}D_i(t)}{\mathrm{d}t} = K(i)D_i(t)$ and get $D_i(t)$.
9. Multiply D_i with the next site tensor B_{i+1} and get the orthogonality center C_{i+1} at the next site.

Higher orders can be obtained by doing step 3 to 6 recursively.