

Tree tensor networks in quantum chemistry

Three-legged tree tensor network states

Ph.D. Thesis

Dissertation submitted in fulfillment of the requirements
for the degree of Doctor (Ph.D.) in Sciences: Physics
Academic year 2020-2021

by

Klaas Gunst

Supervisor: Prof. Dr. Dimitri Van Neck
Co-supervisor: Prof. Dr. Frank Verstraete





Supervisors

Prof. Dr. Dimitri Van Neck	Ghent University
Prof. Dr. Frank Verstraete	Ghent University

Members of the examination committee

Prof. Dr. Patrick Bultinck	Ghent University
Prof. Dr. Stijn De Baerdemacker	University of New Brunswick
Prof. Dr. Jutho Haegeman	Ghent University
Prof. Dr. Örs Legeza	Wigner Research Centre for Physics
Prof. Dr. Jan Ryckebusch	Ghent University
Prof. Dr. Louis Vanduyfhuys	Ghent University

The research described in this thesis was carried out at the Center for Molecular Modeling and the Quantum Group, Ghent University, the Wigner Research Centre for Physics, and at the Division of Chemistry and Chemical Engineering, California Institute of Technology. This research was funded by a Ph.D. fellowship and a travel grant for a long stay abroad from the Research Foundation Flanders (FWO Vlaanderen); and was carried out using the Stevin Supercomputer Infrastructure at Ghent University.



Dankwoord

Hoewel het verwezenlijken van een doctoraat vaak een solitair werk is, hebben veel mensen, direct of indirect, een impact gehad op mijn onderzoek. Ik neem graag eerst de tijd om hen te bedanken voor de afgelopen vier jaar.

Allereerst wil ik mijn promotoren bedanken. Ik wil Dimitri Van Neck bedanken voor het begeleiden van zowel mijn masterthesis als doctoraatsthesis. Reeds als docent wist hij me warm te maken voor de kwantumwereld. Frank Verstraete bedank ik om een antwoord te bieden op de vele vragen aangaande tensornetwerken en voor het aanreiken van een veelheid aan ideeën. Een doctoraat duurt maar vier jaar en dus zijn er onvermijdelijk slechts enkele van zijn vele voorgestelde projecten uitgevoerd.

Zowel mijn collega's op het CMM als in de Sterre zorgden ervoor dat ik het nooit een sleur vond om naar mijn werk te gaan. In het bijzonder wil ik mijn kantoorgenoten door de jaren heen bedanken voor de koffiepauzes en interessante gesprekken (al dan niet over wetenschap) – bedankt Stijn, Pieter, Caitlin en Arthur. Voor het grondig nalezen van dit manuscript ben ik Pieter nog eens extra dankbaar.

Ook wil ik Sebastian bedanken; als begeleider van mijn masterproef heeft hij mij geïntroduceerd tot de wereld van tensornetwerken in de kwantumchemie.

I can only call myself lucky for the hospitality and guidance provided at the Wigner Research Centre for Physics, Budapest, and the California Institute of Technology, Pasadena, during my PhD.

To Örs Legeza, thank you for guiding me around in the world of quantum chemical tensor networks. The many discussions we had during my stay in Budapest were extremely helpful. The patience you showed whilst explaining the nitty-gritty for chemical tensor network algorithms were essential for realizing my first working T3NS code. Later on, your unceasing enthusiasm for your scientific research kept further inspiring me.

To Garnet Chan, thank you for the hospitality and the fruitful discussions during my stay at Caltech. While tree tensor networks never really left my mind, you encouraged me to delve into other quantum chemical methods through several smaller projects.

Both my research stays abroad were made much more pleasant by the people I met there – being it colleagues livening up the office or friends to climb

or relax at the beach with. In particular, I would like to thank Henrik for the close collaboration in Caltech, the hikes in the mountains and proofreading this dissertation.

Verder ook bedankt aan al de vrienden die me lieten ontsnappen aan het wetenschappelijk onderzoek, ware het met een avondje op café, een klimsessie, een trektocht, een ‘mystery weekend’, of een bezoek in Hongarije en een roadtrip in Amerika.

Mijn ouders hebben mij altijd onvermoeibaar gesteund in mijn beslissingen, ook al betekende het dat ik voor enkele maanden naar het buitenland verdween – bedankt voor alles. Thijs heeft mij als oudere broer reeds op jonge leeftijd warm gemaakt voor wetenschap. Ik herinner me nog levendig hoe hij, na er zelf over te leren op school, me introduceerde tot de wereld van atomen en elektronen.

Als laatste wil ik graag Karen bedanken. Ik ben blij dat ik de laatste hordes van mijn doctoraat samen met jou heb kunnen nemen en je me op gepaste tijden het schrijven opzij liet schuiven.

*Klaas Gunst
Gent, 2020*

Contents

Dankwoord v

Samenvatting xi

List of Publications xvii

1 Introduction 1

- 1.1 Entanglement 3
 - 1.1.1 The Schmidt decomposition 4
 - 1.1.2 Entanglement in physical systems 7
- 1.2 The chemical Hamiltonian 9
 - 1.2.1 Finite basis sets and orbitals 10
 - 1.2.2 Second quantization 11
 - 1.2.3 Symmetries 13
- 1.3 Approximate methods 15
- 1.4 A brief overview of this dissertation 17

2 Tensor network states 19

- 2.1 A concise history 20
- 2.2 Modern DMRG 22
 - 2.2.1 The matrix product state 22
 - 2.2.2 Matrix product operators 27
 - 2.2.3 The DMRG algorithm 29
- 2.3 DMRG in quantum chemistry 32
 - 2.3.1 Renormalized operators 33
 - 2.3.2 Resource costs of the algorithm 37
- 2.4 Other tensor networks 38
 - 2.4.1 Tree Tensor Networks 40
- 2.5 The T3NS ansatz 42
 - 2.5.1 The T3NS algorithm 43
 - 2.5.2 Resource costs of the algorithm 44

2.6	Shape of the tree, choice and ordering of the orbitals	45
2.7	Fermionic Networks	45
2.8	Numerical results	46
2.8.1	Lithium fluoride	47
2.8.2	The nitrogen dimer	49
2.8.3	The bisoxo and peroxy isomer of $[\text{Cu}_2\text{O}_2]^{2-}$	50
2.9	A short recap	53
3	Symmetries	55
3.1	A symmetry adapted three-legged tensor	57
3.1.1	Labeling the basis states	57
3.1.2	Reduced tensors	58
3.2	Building the wave function	62
3.2.1	A tensor contraction	62
3.2.2	The canonical form	66
3.2.3	The complete ansatz	70
3.3	Building the chemical Hamiltonian	72
3.4	Optimizing the wave function	77
3.4.1	The two-site tensor	78
3.4.2	The $\text{SU}(2)$ -adapted renormalized operators	79
3.4.3	The effective Hamiltonian: the T3NS case	82
3.4.4	The eigenvalue problem	84
3.4.5	Singular value decomposition with $\text{SU}(2)$ -symmetry	85
3.5	Numerical results	86
3.5.1	The Bisoxo and Peroxy Isomer of $[\text{Cu}_2\text{O}_2]^{2-}$	88
3.6	A short recap	91
4	Seniority	93
4.1	Seniority and tensor networks	95
4.1.1	Suboptimal decomposition	97
4.1.2	DOCI and tensor networks	98
4.2	Applications	99
4.2.1	Nitrogen dimer	100
4.2.2	Benzene	104
4.2.3	Neon dimer	106
4.2.4	A rubidium phenanthrene coordination complex	109
4.3	A short recap	110

5 Orbitals in tensor networks 113

- 5.1 The orbital choice 114
- 5.2 Ordering the orbitals 114
- 5.3 Reordering the orbitals 117
 - 5.3.1 The local permutation 118
 - 5.3.2 Optimizing the orbital ordering 120
 - 5.3.3 Numerical examples 121
- 5.4 Deforming the tree tensor network 126
- 5.5 A short recap 126

6 Conclusions 129**A Two-site optimization 135**

- A.1 The approximation theorem 135
- A.2 One-site versus two-site optimization 137

B Wigner symbols 139

- B.1 The Wigner $3j$ -symbol 140
- B.2 The Wigner $6j$ -symbol 141
- B.3 The Wigner $9j$ -symbol 141

C Some odds and ends for $SU(2)$ 143

- C.1 The QR-decomposition 143
- C.2 The identity 144
- C.3 Some more contractions 145
 - C.3.1 Physical update of a renormalized operator 145
 - C.3.2 Branching update of a renormalized operator 145
 - C.3.3 The effective Hamiltonian: The DMRG case 146

Bibliography 149



Samenvatting

*Voor ik vergeet en later alles anders heet
Voor ik vergeet en ik de feiten en de cijfers
en de namen van de schrijvers niet meer weet*

Spinvis

In 1838 merkte de filosoof Auguste Comte in zijn *Cours de philosophie positive* op dat “elke poging om wiskundige methoden te gebruiken bij de studie van chemische vraagstukken als diep irrationeel en in strijd met de geest van de chemie moet worden beschouwd”. Dertig jaar later leek dit sentiment compleet afwezig toen Dmitri Mendelejev de gekende chemische elementen aan de hand van hun eigenschappen in een tabel ordende. Overtuigd van het periodiek karakter van de elementen loste Mendelejev hiaten in deze tabel op door onder andere het bestaan van scandium, gallium, technetium, germanium en protactinium correct te voorspellen. Ondanks het succes van het periodiek systeem bleef het nog wachten op het ontstaan van de kwantummechanica in het begin van de vorige eeuw voor een verklaring van dit gestructureerde gedrag van de atomen. Kwantummechanica zorgde er samen met het ontstaan van de computer voor dat chemische systemen niet enkel meer in de proefbuis bestudeerd worden. De afwezigheid van wiskunde in de chemie lijkt tegenwoordig ondenkbaar.

Wie geïnteresseerd is in het gedrag van moleculen en dit via computationele technieken wil simuleren, dient een reeks benadering in te voeren om het systeem behapbaar te maken. Enkel heel eenvoudige systemen zoals het waterstofatoom zijn immers exact oplosbaar. Zo abstraheert men vaak de atoomkernen van de molecule tot vaste puntladingen (een gevolg van de Born-Oppenheimer approximatie) zodat enkel de beweging van de elektronen uitgerekend hoeft te worden. Relativistische effecten worden ook vaak verwaarloosd door aan te nemen dat de elektronen veel trager bewegen dan de lichtsnelheid. Het beperken van de bewegingsruimte van de elektronen is een derde, vaak noodzakelijke, benadering. De bewegingsvrijheid van elk elektron wordt beperkt tot een aantal *orbitalen* in plaats van de complete driedimensionale ruimte. Deze orbitalen worden vaak gekozen zodat ze reeds de karakteristieke

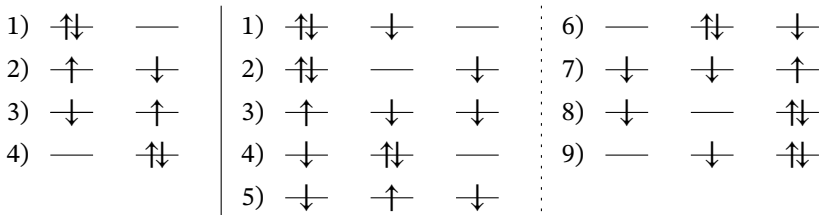
eigenschappen van de elektronenbeweging rond een atoom vertonen. Op deze manier kunnen we het aantal nodige orbitalen voor een accurate beschrijving van de molecule beperkt houden. Bij elk van deze benaderingen moeten we stilstaan bij hun geldigheid. De opgenoemde benaderingen zijn niet universeel geldig, zo is het verwaarlozen van relativistische effecten bijvoorbeeld niet verantwoord voor de studie van moleculen met zware atomen. In deze thesis zullen we echter steeds gebruik maken van de voorgenoemde benaderingen en enkel moleculen bestuderen waar deze geldig zijn.

Elektronen zijn fundamentele deeltjes met halve *spin*. Concreet betekent dit dat een elektron een intrinsieke eigenschap bezit die de waarde *op*, *neer*, of een *superpositie* van op en neer kan zijn. Deze toestanden worden wiskundig als volgt weergegeven:

$$\begin{array}{ll} |\uparrow\rangle & \text{(op)} \\ |\downarrow\rangle & \text{(neer)} \\ c_{\uparrow}|\uparrow\rangle + c_{\downarrow}|\downarrow\rangle & \text{(superpositie)} \end{array}$$

waar c_{\uparrow} en c_{\downarrow} complexe getallen zijn. De superpositie van op en neer betekent dat een meting van deze spin op of neer als resultaat zal geven met een respectievelijke kans van $|c_{\uparrow}|^2$ en $|c_{\downarrow}|^2$. Het bestaan van superposities is een cruciale eigenschap van de kwantummechanica, en maakt metingen op fysische systemen intrinsiek probabilistisch.

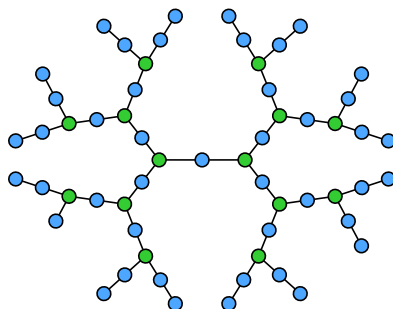
Elektronen zijn fermionen en worden gekarakteriseerd door hun orbitaal en hun spintoestand. Elektronen volgen, net als alle fermionen, het uitsluitingsprincipe van Pauli; twee elektronen kunnen niet tegelijk in eenzelfde toestand verkeren. Er kunnen dus geen twee elektronen zowel dezelfde ruimtelijke orbitaal bezetten als dezelfde spin hebben.



Figuur 1: (links) de mogelijke configuraties voor twee elektronen (een met spin op, een met spin neer) in een systeem met twee ruimtelijke orbitalen. (rechts) de mogelijke configuraties voor drie elektronen (een met spin op, twee met spin neer) in een systeem met drie ruimtelijke orbitalen. De orbitalen worden met horizontale strepen weergegeven. Elke orbitaal kan leeg, enkelvoudig (\uparrow of \downarrow), of dubbel ($\uparrow\downarrow$) bezet zijn.

Laten we nu de oplossingsruimte voor zulke kwantummechanische syste-

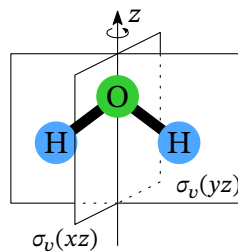
men onder de loep houden, rekening houdend met bovenstaande bemerkingen. In figuur 1 worden de vier mogelijke configuraties voor twee elektronen (één \uparrow , één \downarrow) in twee orbitalen en de negen mogelijke configuraties voor drie elektronen (één \uparrow , twee \downarrow) gegeven. Het toevoegen van een elektron en een orbitaal heeft het aantal mogelijke configuraties van ons systeem meer dan verdubbeld. Bij het toevoegen van nog een extra \uparrow elektron en een ruimtelijke orbitaal zwellen de configuraties aan tot 36. Deze sterke groei van het aantal configuraties in functie van de systeemgrootte is de typerende *exponentiële muur* van de kwantummechanica. Het neemt danige proporties aan dat het bestuderen van tien \uparrow en tien \downarrow elektronen in twintig orbitalen reeds miljarden configuraties vraagt en dit systeem verdubbelen qua grootte verandert de 10^9 mogelijke configuraties in 10^{22} configuraties. De onhoudbaarheid van deze situatie is vrij duidelijk, zeker aangezien het systeem in een superpositie van al deze configuraties tegelijk kan verkeren. Voor het kwantummechanisch bestuderen van moleculen hebben we dus nood aan verdere benaderingen. Enkel voor de allerkleinste moleculen is het zoeken naar de grondtoestand (de energetisch meest gunstige toestand) in de exponentieel groeiende oplossingsruimte mogelijk; voor ietwat grotere systemen wordt het al snel zoeken naar een speld in een hooiberg.



Figuur 2: Een grafische voorstelling van een ‘driebenige boomvormige tensornetwerktoestand’ (T3NS). De cirkels representeren *tensoren* die verbonden zijn met elkaar in een netwerk. De blauwe cirkels zijn tensoren geassocieerd met orbitalen van het systeem, terwijl groene cirkels tensoren zijn die louter dienen voor het *verstrengelen* van verschillende orbitalen; ze verbinden de verschillende takken van het boomvormige netwerk. Dit netwerk wordt boomvormig genoemd omdat het vertakkingen maar geen lussen vertoont.

Er bestaan veel mogelijke manieren om de exponentieel groeiende oplossingsruimte verder te benaderen, maar niet elke benadering is even accuraat of even wijd toepasbaar. In deze thesis bestuderen we de zogenaamde ‘driebenige boomvormige tensornetwerktoestanden’ (de three-legged tree tensor network state of T3NS). Een grafische voorstelling van een T3NS wordt gegeven

in figuur 2. Het is een methode uit de overkoepelende klasse van de tensor-netwerkt toestanden. Bij deze toestanden worden zogenaamde tensoren (een veralgemening van matrices naar hogere dimensie) toegewezen aan de verschillende orbitalen. Door deze tensoren te verbinden met elkaar komen de orbitalen met elkaar in *kwantumverstrengeling*, dit zijn kwantummechanische correlaties die ontstaan door een superpositie van toestanden. Tensornetwerken zijn voornamelijk efficiënt voor het bestuderen van systemen met een lage kwantumverstrengeling. Ze maken het numeriek oplossen van grote moleculen mogelijk door niet de hele oplossingsruimte af te zoeken voor de grondtoestand, maar zich te beperken tot enkel toestanden met een lage kwantumverstrengeling. Gelukkig lijken de meeste systemen in de natuur inderdaad een lage verstrengeling te vertonen. In hoofdstuk 2 geven we een overzicht van tensornetwerken die gebruikt worden in de kwantumchemie en introduceren we de T3NS. We bespreken ook hoe deze netwerken kunnen gebruikt worden voor het vinden van moleculaire grondtoestanden en tonen enkele initiële berekeningen.



Figuur 3: Schematische voorstelling van een watermolecule. Deze molecule is invariant onder rotaties van 180° rond de z -as en is ook invariant voor reflecties langs het xz - en yz -vlak.

Een andere manier om het probleem te vereenvoudigen is door rekening te houden met de symmetrieën van het systeem. Moleculaire systemen vertonen verscheidene symmetrieën. Als we bijvoorbeeld naar een watermolecule in figuur 3 kijken, merken we al snel dat deze niet verandert onder rotaties van 180° rond de z -as of wanneer we reflecteren langs het xz - of yz -vlak. Sommige symmetrieën zijn minder voor de hand liggend. Zo staat het feit dat er geen elektronen gecreëerd of vernietigd worden in de molecule ook in verband met een symmetrie. Herinnert u zich ook nog de spintoestand van het elektron dat zowel op als neer kan zijn. Het begrip ‘op’ en ‘neer’ is echter afhankelijk van de gekozen meetrichting van de spin en er is natuurlijk geen enkele reden waarom een specifieke meetrichting de voorkeur zou genieten. Elke andere richting is een even geldige keuze voor het meten van de elektronspin. Deze keuzevrijheid zorgt voor nog een symmetrie aanwezig in moleculen. Door

deze symmetrieën kunnen enkele voorwaarden voor de grondtoestand van de molecule geformuleerd worden, wat de zoektocht ernaar vereenvoudigt. Het gebruik van de symmetrieën van de molecule voor het formuleren van een efficiënter T3NS algoritme is het onderwerp van hoofdstuk 3.

De technieken die ontwikkeld worden in hoofdstuk 3 kunnen op hun beurt verder gebruikt worden om extra voorwaarden aan de grondtoestand op te leggen. Hoewel de exacte grondtoestand niet noodzakelijkerwijs voldoet aan deze voorwaarden, kan het ons toch verder helpen indien de best benaderende toestand die wel voldoet accuraat genoeg is. In figuur 1 toonden we orbitalen bezet door geen, één, of een paar elektronen. In de molecule is er geen symmetrie aanwezig die ons meer leert over het aantal enkelbezette orbitalen. Toch kan het beperken van het aantal ongepaarde elektronen accurate resultaten opleveren. De studie van het vinden van de grondtoestand met slechts een beperkt aantal enkelbezette orbitalen is het onderwerp van hoofdstuk 4.

In de T3NS worden de verschillende orbitalen verstrengeld door hun tensoren met elkaar te verbinden in een netwerk zoals grafisch weergegeven in figuur 2. Een groep orbitalen die in hoge mate verstrengeld zijn met elkaar worden best dicht bij elkaar gezet in het netwerk. Het beïnvloedt de efficiëntie van het algoritme op een gunstige manier. Welke orbitalen sterk verstrengeld zijn met elkaar in de grondtoestand is echter niet op voorhand geweten. In hoofdstuk 5 bespreken we methoden voor het bepalen van een initiële ordening van de orbitalen en het verder optimaliseren van de ordening op het netwerk tijdens de uitvoering van het T3NS algoritme.



List of Publications

- [1] K. Gunst, S. Wouters, S. De Baerdemacker, and D. Van Neck, “Block product density matrix embedding theory for strongly correlated spin systems”, *Physical Review B* **95**, 195127 (2017).
- [2] K. Gunst, F. Verstraete, S. Wouters, Ö. Legeza, and D. Van Neck, “T3NS: three-legged tree tensor network states”, *Journal of Chemical Theory and Computation* **14**, ACS Editors’ Choice, 2026–2033 (2018).
- [3] C. Lanssens, P. W. Ayers, D. Van Neck, S. De Baerdemacker, K. Gunst, and P. Bultinck, “Method for making 2-electron response reduced density matrices approximately N-representable”, *The Journal of Chemical Physics* **148**, 084104 (2018).
- [4] K. Gunst, F. Verstraete, and D. Van Neck, “Three-legged tree tensor networks with SU(2) and molecular point group symmetry”, *Journal of Chemical Theory and Computation* **15**, 2996–3007 (2019).
- [5] K. Gunst, D. Van Neck, P. A. Limacher, and S. De Baerdemacker, “The seniority quantum number in tensor network states”, arXiv e-prints, 2008.00733 (2020).



1

Introduction

*You will have to concede, Herr Pauli,
by partially waiving your exclusion principle
you might free us from many worries of daily life,
for instance from the traffic problem in our streets.*

P. Ehrenfest [1]

This introduction sets the stage for our study of tensor networks in quantum chemistry. In Section 1.1, we give a brief introduction to entanglement. This counterintuitive quantum mechanical correlation lies at the source of the recent success of tensor network methods; a proper understanding of this phenomenon is welcome and will be addressed.

In Section 1.2, we discuss the quantum chemical Hamiltonian. The study of the electronic structure in molecules is central in this dissertation; as such we need to understand the approximations made and which useful properties this Hamiltonian exhibits.

In Section 1.3, a very brief overview of some well-established electronic structure methods are given. We also discuss strong and weak correlations, the two major types of electronic correlations distinguished by chemists. Most methods are only successful in retrieving one type of correlation, while methods that are potent enough to retrieve both are often computationally very expensive.

We conclude the chapter with an overview of the remainder of this dissertation.

The nineteenth-century philosopher Auguste Comte remarked in his *Cours de Philosophie Positive* that “every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry” [2]. In contrast, scientists now rely more and more on computational insight for the understanding of physical and chemical phenomena in molecules. The advent of quantum mechanics in the first half of the twentieth century is, without any doubt, the major factor for this sudden shift. It gave physicists and chemists a theoretical framework for the study of systems on the atomic scale; their behavior is governed by the Schrödinger equation¹ [3], which is probably one of the most famous equations in physics and rightly so. It is given by

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle \quad (1.1)$$

or for time-independent systems

$$\hat{H} |\Psi\rangle = E |\Psi\rangle. \quad (1.2)$$

The *Hamiltonian* of the system is represented by \hat{H} . It is an operator containing information on the kinetic energy of the particles and the interactions between them. $|\Psi\rangle$ represents the state of the system; it reflects the inherent probabilistic nature of quantum mechanics as multiple states can occur in *superposition*. To further exemplify this, let us assume we perform a quantum mechanical coin toss. The result of the coin toss can be either heads or tails; let us represent these outcomes by the states $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively. Before the outcome of the coin toss is measured we could represent the state of the coin in a superposition of heads and tails, i.e.

$$|\Psi\rangle = C_{\uparrow} |\uparrow\rangle + C_{\downarrow} |\downarrow\rangle. \quad (1.3)$$

The exact outcome of the measurement is *a priori* unknown; we only know the probabilities of measuring heads and tails to be $|C_{\uparrow}|^2$ and $|C_{\downarrow}|^2$ respectively. According to the Copenhagen interpretation of quantum mechanics, the wave function of our coin will collapse to either $|\uparrow\rangle$ or $|\downarrow\rangle$ after measurement; we are now completely certain about the state of our coin.

Let us now continue to the toss of two distinctive coins. Instead of two outcomes we have four possible outcomes which we will again represent by $|\uparrow\uparrow\rangle$, $|\downarrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$ and $|\downarrow\downarrow\rangle$. The wave function (or state) of the system is now given by

$$|\Psi\rangle = C_{\uparrow\uparrow} |\uparrow\uparrow\rangle + C_{\downarrow\uparrow} |\downarrow\uparrow\rangle + C_{\uparrow\downarrow} |\uparrow\downarrow\rangle + C_{\downarrow\downarrow} |\downarrow\downarrow\rangle. \quad (1.4)$$

¹At least, in the non-relativistic case.

The probabilities of the different outcomes are again given by the squared value of their amplitudes C . Measuring even one coin will result into an appropriate collapse of the total wave function of our system. For example, when measuring the outcome to be heads for our first coin the wave function will collapse to $|\Psi'\rangle = C_{\uparrow\uparrow}|\uparrow\uparrow\rangle + C_{\uparrow\downarrow}|\uparrow\downarrow\rangle$ up to normalization.

Adding one extra coin doubles the number of possible outcomes and the number of needed amplitudes. This trend will continue with each added coin; the toss of N coins will need 2^N different amplitudes for 2^N different outcomes. This exponential growth of the *Hilbert space* for increasing system sizes is the biggest hurdle for anyone practicing many-body quantum physics.

1.1. Entanglement

As many analogies, the coin toss is not airtight. One could easily raise the objection that the coin toss is not inherently probabilistic; it is still possible to predict the outcome of the toss before the measurement, if only the conditions of our problem were known up to arbitrary accuracies. It is through a lack of knowledge that we perceive the coin toss as probabilistic. Couldn't we raise the same objection about quantum mechanics?

Albert Einstein believed exactly this. As an advocate of the hidden variable theory, he believed that quantum mechanics provided an incomplete description of reality; there are undiscovered variables that govern the exact outcome of quantum measurements. In 1935 he published together with Boris Podolsky and Nathan Rosen the famous EPR-paradox [4]. While this paradox was originally formulated as a measurement of position and momentum of two particles, Bohm later reformulated it for another set of non-commuting operators. We will reiterate Bohm's more practical realization of the EPR-paradox [5]. Imagine two spin-half particles A and B which couple to a singlet. The quantum mechanical wave function of this system is given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}}[|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B] \quad (1.5)$$

where we have neglected all degrees of freedom but the spin. The states $|\uparrow\rangle$ and $|\downarrow\rangle$ refer to the spin projection in a certain direction; let's say the z -direction. As a next step, let us imagine these particles are sent far away in opposite directions; although they still form a singlet we assume they are too far to directly interact with each other. When performing a spin measurement on particle A along the z -axis, we will, with full certainty, obtain the opposite result for measurements along the same axis for particle B ; the measurement on particle A has made the S_z property of particle B definite due to the wave function collapse. On the other hand, measurements of the spin along the x -axis of B

will still be probabilistic, there is fifty-fifty chance for either spin up or spin down. As quantum mechanics is probabilistic, this poses a problem according to the EPR-paradox. While the spin of B is inherently indefinite in all directions before measurements on A , it becomes suddenly definite in the same direction of the measurement on A , but only in this direction. If we assume that the inherent probabilistic nature of quantum theory is correct, there has to be some instantaneous interaction between A and B ² that communicates the measuring direction on A . EPR concluded that the quantum theory cannot be complete as they excluded the existence of any ‘spooky action at a distance’ [6]. A more complete hidden variable theory had to be waiting around the corner.

David Bohm successfully formulated a hidden variable theory in correspondence with quantum mechanics by extending the pilot wave theory of Louis de Broglie [7–9]. However, Bohmian mechanics also has its controversies. Although deterministic, it is manifestly *nonlocal* when entangled³ particles are studied. It enticed John Stewart Bell to consider if any local hidden variable theory that corresponds with the quantum theory is at all possible. His pondering resulted into *Bell’s theorem* [10]; it provides an inequality respected by any local hidden variable theory but violated by quantum mechanics. Multiple experimental measurements suggest that nature indeed violates this inequality; local hidden variable theories cannot adequately represent our reality [11–15].

Whether nature’s spooky interaction at a distance is due to nonlocalities as in Bohmian mechanics, through wave function collapse as in the Copenhagen interpretation or through any of the numerous other interpretations of quantum mechanics, is out of scope for this dissertation. Only the fact that entanglement exists is of importance for the further understanding of this dissertation. We will conclude this discussion by quoting one of the greats of the past century:

The interpretation of quantum mechanics has been dealt with by many authors, and I do not want to discuss it here. I want to deal with more fundamental things.

P.A.M. Dirac

1.1.1. The Schmidt decomposition

Let us unveil the spookiness a bit by mathematically defining entanglement. First, we need to clarify the existence of two types of states in quantum mechanics, i.e. the *pure states* and *mixed states*. Pure states are any state of a

²Or between the measuring device on A and the particle B .

³The modern term for the spooky action at a distance.

system that can be represented by a ket $|\Psi\rangle$, i.e. a single vector in the Hilbert space H . The density matrix of a pure state is given by $\rho = |\Psi\rangle\langle\Psi|$. Mixed states, on the other hand, cannot be represented by a single wave function; they can only be represented by their density matrix ρ which is a positive semi-definite operator with trace one. By using the spectral decomposition of ρ , given by

$$\rho = \sum_{i=1}^{\dim(H)} \eta_i |\Psi_i\rangle\langle\Psi_i|, \quad (1.6)$$

we can realize that mixed states are convex combinations of pure states; the density matrix represents a statistical ensemble of pure states. Mixed states are useful when the system studied is not completely isolated from its environment.

Entanglement represents the quantum mechanical correlations between two subsystems of the complete isolated system. For example the wave function discussed in Bohm's variation on the EPR-paradox in Eq. (1.5) assumes a partitioning of the two-particle Hilbert space into the single-particle Hilbert spaces of A and B . In quantum chemistry for example, the partitioning will often be defined as the Fock spaces spanned by a subset of the orbitals in the system. Once we have made up our mind on how to partition the system the Schmidt decomposition proves to be an indispensable tool [16]. The Schmidt decomposition states the following:

Theorem 1.1 (Schmidt decomposition) *Let us assume the pure wave function $|\Psi\rangle \in H$ where H is the total Hilbert space of the system. For any subspaces H_A and H_B for which $H = H_A \otimes H_B$, there exists orthonormal sets $\{|\phi_1\rangle, \dots, |\phi_M\rangle\}$ in H_A and $\{|\chi_1\rangle, \dots, |\chi_M\rangle\}$ in H_B where $M = \min(\dim(H_A), \dim(H_B))$ such that*

$$|\Psi\rangle = \sum_{i=1}^M s_i |\phi_i\rangle |\chi_i\rangle \quad (1.7)$$

where s_i are real, non-negative numbers uniquely defined for $|\Psi\rangle$.

The Schmidt decomposition is actually a reformulation of the singular value decomposition (SVD) for vectors with respect to their Hilbert space. It provides the minimal amount of orthonormal states needed for an exact representation (when disregarding states connected to Schmidt values of zero) and the best approximation of the total wave function by smaller sets of orthonormal states is given by a truncated Schmidt decomposition. For more information we refer the reader to Appendix A.

One can easily see that subsystems A and B with Hilbert spaces H_A and H_B are more entangled (i.e. are more correlated) if there are a lot of large Schmidt

values in Eq. (1.7); when doing subsequent measurements on subsystems A and B , the result on B will be correlated with the outcome of the measurement on A . On the other hand, when there is only one non-zero Schmidt value we have a *separable state*; any measurement on A will not influence measurements on B . In the EPR-paradox the state in Eq. (1.5) is *maximally entangled* as a maximum amount of Schmidt states are needed and all Schmidt values are of equal magnitude, i.e. $1/\sqrt{M}$ for a normalized pure state.

The Schmidt values give us a means to express the entanglement of a wave function between two subsystems. However, instead of a vector of singular values it would be nice to express entanglement also as a single scalar. Von Neumann introduced an entropy based on the classical Shannon entropy. It is suitably called the von Neumann entropy [17] and is given by the Shannon entropy of the spectrum of ρ , i.e. the density matrix of the mixed state. The von Neumann entropy is given by

$$S(\rho) = -\text{tr}(\rho \ln \rho) = -\sum_{i=1}^{\dim(H)} \eta_i \ln \eta_i \quad (1.8)$$

where η_i are the eigenvalues of ρ . For any pure state we immediately notice that the von Neumann entropy is zero as $\rho = |\Psi\rangle\langle\Psi|$ and thus $\eta = (1, 0, \dots, 0)$.

To calculate the entropy of subsystem A of the pure state in Eq. (1.7), the reduced density matrix of the state limited to subsystem A has to be obtained. This is done by a partial trace of the full density matrix over subsystem B , i.e.

$$\rho_A = \text{tr}_B(\rho) = \text{tr}_B\left(\sum_{i=1}^M s_i |\phi_i\rangle\langle\chi_i|\right) \sum_{j=1}^M s_j \langle\phi_j|\langle\chi_j| = \sum_{i=1}^M s_i^2 |\phi_i\rangle\langle\phi_i|. \quad (1.9)$$

We immediately obtain a spectral decomposition of ρ_A ; the eigenvalues are given by the square of the Schmidt values while the eigenstates are given by the Schmidt orthonormal states. By taking the partial trace of the pure state over subsystem B , we have obtained a reduced density matrix corresponding to a mixed state for subsystem A . Now, the von Neumann entropy of subsystem A is given by

$$S(\rho_A) = -\sum_{i=1}^M s_i^2 \ln s_i^2. \quad (1.10)$$

It is easy to check that a separable state for a given partition results into a von Neumann entropy of zero while a maximally entangled state results into the maximal possible von Neumann entropy $S(\rho_A) = \ln M$. The same analysis holds for subsystem B , i.e. $S(\rho_A) = S(\rho_B)$ and the eigenvalues and eigenstates of ρ_B are given by the square of the Schmidt values and the Schmidt states $|\chi_i\rangle$.

The von Neumann entropy is not the only possible entanglement entropy measure. Another option is the Rényi entropy [18, 19] given by

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \ln(\text{tr } \rho^\alpha) = \frac{1}{1-\alpha} \ln\left(\sum_{i=1}^M \eta_i^\alpha\right) \quad (1.11)$$

for $0 \leq \alpha < 1$. For the limit $\alpha \rightarrow 1$ the von Neumann entropy is obtained. Just as for the von Neumann entropy, any separable state has a zero entropy for the Rényi entropy while the maximal value of $S_\alpha = \ln M$ is obtained for any maximally entangled state.

When measuring the entanglement entropy, small α -values of the Rényi entropy will give more weight to the smallest Schmidt values; slowly decaying tails in the Schmidt values will be penalized heavier.

1.1.2. Entanglement in physical systems

Let us briefly discuss how entanglement behaves in physical systems. We will start by investigating the low-lying eigenstates of the Heisenberg Hamiltonian; we consider a one-dimensional chain of twenty $S = 1/2$ spins and impose periodic boundary conditions. The Hamiltonian of the system is given by

$$\hat{H} = J \sum_{\langle ij \rangle} \hat{S}_i \cdot \hat{S}_j, \quad (1.12)$$

where $\langle ij \rangle$ denotes the neighboring spins; it is a system with only local interactions.

The local Hilbert space of each spin is given by \mathbb{C}^2 ; each spin can either be spin-up, spin-down or any linear combination. The total Hilbert space is a direct product of the local Hilbert spaces; it is given by $(\mathbb{C}^2)^{\otimes N}$ with N the number of spins. We can easily sample this space by normalizing vectors whose elements are drawn as independent and identically normal distributed random variables⁴. For these random wave functions we can easily calculate the von Neumann entropy for segments of different sizes A . In Fig. 1.1 the average von Neumann entropy for different segment sizes A are given for 100 000 samples. The variances are also shown as error bars, albeit barely visible. As could be expected, the von Neumann entropy increases linearly with A . When comparing the random states with the lowest lying eigenstates of the Heisenberg Hamiltonian, it is quite apparent that these behave very differently. Again looking

⁴Indeed, the probability density for a given vector $\vec{x} = (x_1, \dots, x_n)$ is given by $p(x_1, \dots, x_n) = \prod_n \mathcal{N}(0, 1) dx_i \propto e^{-\frac{1}{2} \|\vec{x}\|^2} dx_1 \dots dx_n$. The probability of drawing a vector is only dependent on its magnitude and not its direction; every vector in the Hilbert space has equal probability after normalization.

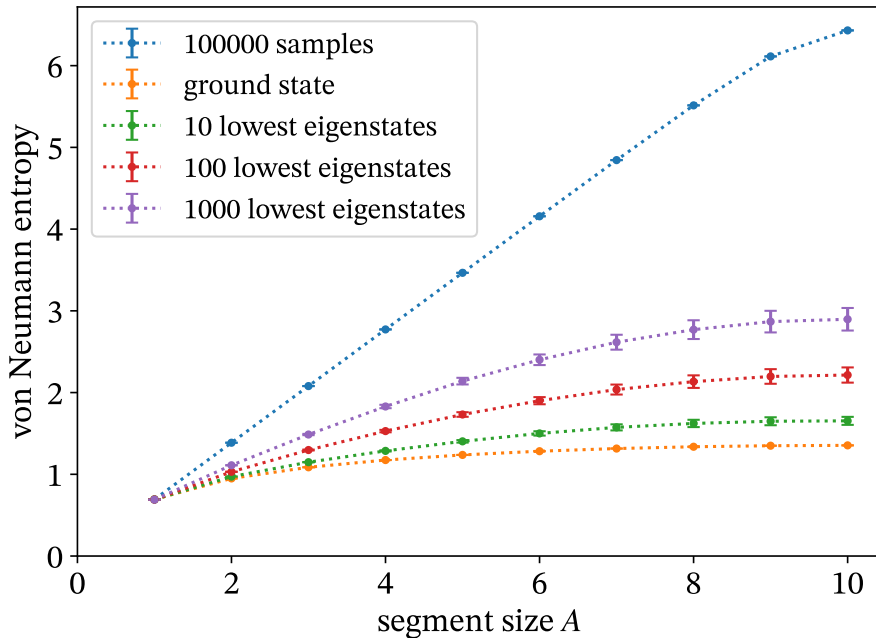


Figure 1.1: Illustration of the area law on a one-dimensional Heisenberg spin chain of twenty $S = 1/2$ spins. Periodic boundary conditions are imposed. The average von Neumann entropy for 100 000 random wave functions is compared with the average von Neumann entropy of the lowest lying eigenstates. Variances are also shown as error bars.

at Fig. 1.1, these eigenstates seem to express an entropy that saturates with subsystem size.

Although this example is quite anecdotal, it is a trend persisting in other systems and system sizes as well; it is known as the *area law* of entanglement. Rather than scaling with the volume A , the entanglement scales with the area of the boundary of subsystem ∂A for ground states in a lot of physical systems. It is as if only the boundary of the subsystem contributes to the quantum mechanical correlations with its environment, which seems a plausible assumption for local interactions. For ground states of gapped local Hamiltonians in one dimension the area law is proven [20]; for general systems in higher dimensions it remains a conjecture [21–23] and for gapless Hamiltonians logarithmic corrections on the area law are needed [22, 24].

In contrast to classical entropy, the von Neumann entropy does not seem to behave extensively. It gives an explanation to the non-intuitive and (initially) controversial character of entanglement as its direct effects on the macroscopic

level are normally minimal. Many systems in nature seem to exhibit ground state wave functions with a low entanglement; only a small corner of the complete Hilbert space is relevant. Tensor network states, the main subject of this dissertation, exploit exactly this tendency for low entanglement by providing an efficient parametrization of the relevant corner of the Hilbert space for systems respecting area laws. This type of ansatz will be introduced in the next chapter.

A much more thorough discussion of entanglement and its implications can be found in *Quantum computation and quantum information* by Nielsen and Chuang [25].

1.2. The chemical Hamiltonian

In this dissertation we will occupy ourselves with the ab initio study of chemical systems. More particularly we will restrict ourselves to the time-independent non-relativistic Schrödinger equation Eq. (1.2). As we are mainly interested in ground states of static systems time-dependency is not needed. Negligence of relativistic effects in studying chemical systems is generally not a problem when restricting ourselves to light atoms. It allows us to work with a fixed particle number; no electrons are created or annihilated as in quantum field theory. Effects arising from the relativistic Dirac equation [26, 27] like velocity-dependent masses, the spin-orbit coupling and the Darwin term are neglected although some relativistic effects can still be taken into account in non-relativistic algorithms by various approximate methods [28–33].

We need to solve the Schrödinger equation in Eq. (1.2) where the exact non-relativistic Hamiltonian is given by

$$\hat{H} = -\sum_i \frac{1}{2} \nabla_i^2 - \sum_A \frac{1}{2M_A} \nabla_A^2 + \frac{1}{2} \sum_{A \neq B} \frac{Z_A Z_B}{\|\vec{R}_A - \vec{R}_B\|} - \sum_{i,A} \frac{Z_A}{\|\vec{R}_A - \vec{r}_i\|} + \frac{1}{2} \sum_{i \neq j} \frac{1}{\|\vec{r}_i - \vec{r}_j\|}, \quad (1.13)$$

where i and j indicate electrons and A and B indicate nucleons with a nuclear charge Z_A and Z_B . The first line consists of the kinetic terms for the nuclei and electrons while the second line sums the different Coulomb interactions. These are repulsive for electrons i and j or nuclei A and B and attractive between electrons and nuclei. Atomic units were used for this Hamiltonian [34]; we choose the electronic charge e and mass m_e , the reduced Planck constant \hbar and the inverse Coulomb constant $4\pi\epsilon_0$ equal to 1.

Next, we will consider the Born-Oppenheimer approximation [35] as the Hamiltonian in Eq. (1.13) is still quite unruly. In this approximation, the mo-

tion of the nuclei and electrons are treated separately. The Born-Oppenheimer approximation is widely used in quantum chemistry and is valid when the potential energy surfaces (PESs) of the different energy levels are well separated. As we are interested in the electronic structure of molecules, we fix the nuclear positions \vec{R}_A and treat them as classical point charges. The simplified Hamiltonian dictating the electronic structure is now given by

$$\hat{H}(\vec{\mathbf{R}}) = E_0(\vec{\mathbf{R}}) - \sum_i \frac{1}{2} \nabla_i^2 - \sum_i V(\vec{\mathbf{R}}, \vec{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{\|\vec{r}_i - \vec{r}_j\|}. \quad (1.14)$$

The nuclear positions are collectively denoted by $\vec{\mathbf{R}}$. By treating the nuclei as classical point charges their Coulomb interactions result in a constant term $E_0(\vec{\mathbf{R}})$ and an electrostatic potential for the electrons $V(\vec{\mathbf{R}}, \vec{r}_i)$.

1.2.1. Finite basis sets and orbitals

The chemical Hamiltonian established in Eq. (1.14) is, although much simplified, still quite unwieldy. At this stage, the Schrödinger equation for the electronic wave function is given by

$$\hat{H}(\vec{\mathbf{R}})|\Psi_e(\vec{r} | \vec{\mathbf{R}})\rangle = E(\vec{\mathbf{R}})|\Psi_e(\vec{r} | \vec{\mathbf{R}})\rangle. \quad (1.15)$$

In general, the partial differential equations (PDEs) one obtains by solving this Schrödinger equation are not exactly soluble. Only the most simple atoms and molecules can be exactly solved; for example, for the hydrogen-like atoms a closed solution exists [3] and the Hylleraas expansion provides a solution for helium-like atoms up to arbitrary precision [36–39]. For general molecular systems further approximations have to be considered.

Traditional methods for numerical solving PDEs, such as finite element methods [40, 41] and finite difference methods [42–46], have been used in quantum chemistry, although they are in the minority. Instead of discretizing, it is much more common to approximate the full three-dimensional single-particle space by an expansion of orbital-like basis functions centered around the different atoms of the system. When expressing the different interactions in any basis set – discretized, Gaussian-type orbitals (GTO) [47–49], Slater-type orbitals (STO) [50], or any other – we need to evaluate the one- and two-electron integrals given by

$$T_{ij} = \int d^3\vec{r}_1 \phi_i^*(\vec{r}_1) \left[\frac{1}{2} \nabla^2 - \sum_A \frac{Z_A}{\|\vec{R}_A - \vec{r}_1\|} \right] \phi_j(\vec{r}_1) \quad (1.16)$$

$$V_{ijkl} = \iint d^3\vec{r}_1 d^3\vec{r}_2 \phi_i^*(\vec{r}_1) \phi_j^*(\vec{r}_2) \frac{1}{\|\vec{r}_1 - \vec{r}_2\|} \phi_k(\vec{r}_1) \phi_l(\vec{r}_2), \quad (1.17)$$

where $\phi_i(\vec{r})$ are the different spatial basis functions. While these integrals are rather sparse for discretized space, they become dense for the non-local GTO and STO basis sets. However, GTOs and STOs are generally preferred as much less basis functions are needed for an accurate description of the electronic structure. While STOs already include a correct radial dependency of $\exp(-\zeta r)$ for the orbitals, GTOs do not; especially the cusp observed at the nuclei is badly represented by the $\exp(-\alpha r^2)$ dependency of the GTOs. One would thus assume STOs are the golden standard for any quantum chemist. Unfortunately, calculating the integrals for STOs becomes very cumbersome when multiple nuclei are involved, while there exist analytic expressions for the calculation of the integrals for GTOs [47]. As such, GTOs are much more widely used.

In this dissertation, we will restrict ourselves to approximating the three-dimensional single-particle space in a set of orthonormal orbitals. Orbitals occupied by electrons with a different spin-projection are orthogonal irrespective of their spatial form. One could choose different sets of spatial orbitals for the up and down electrons. However, we will choose one set of spatial orthogonal orbitals for both electrons as this facilitates the usage of the total spin symmetry. Whether the used spatial orbitals are linear combinations of GTOs, STOs or any other set of basis functions and whether these orbitals were generated by a Hartree-Fock calculation [34, 51–53], a localization method [54–56] or any other orthogonalization procedure does not really matter for the formulation of the algorithm presented in this dissertation. It is of course of great importance for the accuracy and speed of the calculations as tensor network methods are not orbital invariant.

We briefly reiterate the used approximations for studying a static molecule:

- Relativistic effects are neglected; as such we need to solve the Schrödinger equation instead of the Dirac equation.
- Using the Born-Oppenheimer approximation, we model the nuclei as fixed point charges. We are only concerned with the movement of the electrons.
- The infinite three-dimensional space is approximated by using a finite set of basis functions, usually Gaussian-type orbitals.

1.2.2. Second quantization

As a last step, we will write the chemical Hamiltonian of Eq. (1.14) projected onto the orthonormal set of orbitals in its second quantized form [57, 58]:

$$H = E_0 + \sum_{ij} T_{ij} \sum_{\sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + \frac{1}{2} \sum_{ijkl} V_{ijkl} \sum_{\sigma\tau} c_{i\sigma}^{\dagger} c_{j\tau}^{\dagger} c_{l\tau} c_{k\sigma}. \quad (1.18)$$

The introduced operators $c_{i\sigma}^\dagger$ and $c_{i\sigma}$ are the fermionic creation and annihilation operators, respectively. They create (and annihilate) electrons in the corresponding spin-orbital. The indexes i, j, k and l denote the spatial orbital, while σ and τ are the spin-projections of the electron ($\sigma, \tau \in \{\uparrow, \downarrow\}$). States can be created by applying a string of second quantized creation operators onto the vacuum. For example, a state of a single up-electron in orbital i is represented by

$$|\phi_{i\uparrow}\rangle = c_{i\uparrow}^\dagger|-\rangle, \quad (1.19)$$

while a general many-body state in L orbitals is given by a linear combination of strings of creation operators, i.e.

$$\begin{aligned} |\Psi\rangle &= \sum_{\{n_{i\sigma}\}} C_{\{n_{i\sigma}\}} |n_{1\uparrow} n_{1\downarrow} \dots n_{L\uparrow} n_{L\downarrow}\rangle \\ &= \sum_{\{n_{i\sigma}\}} C_{\{n_{i\sigma}\}} (c_{1\uparrow}^\dagger)^{n_{1\uparrow}} (c_{1\downarrow}^\dagger)^{n_{1\downarrow}} \dots (c_{L\uparrow}^\dagger)^{n_{L\uparrow}} (c_{L\downarrow}^\dagger)^{n_{L\downarrow}} |-\rangle, \end{aligned} \quad (1.20)$$

where each $n_{i\sigma} \in \{0, 1\}$. On the other hand, acting upon the vacuum with any annihilation operator destroys the state completely, i.e.

$$c_{i\sigma}|-\rangle = 0 \quad (1.21)$$

Interactions are seen as the subsequent annihilation and creation of electrons. For example, the scattering of two electrons through the Coulomb interaction given by $V_{ijkl} c_{i\sigma}^\dagger c_{j\tau}^\dagger c_{l\tau} c_{k\sigma}$ is viewed as the annihilation of the two electrons in their initial orbitals $\phi_{k\sigma}$ and $\phi_{l\tau}$ followed by the creation of two electrons in the final orbitals $\phi_{i\sigma}$ and $\phi_{j\tau}$.

Fermionic creation and annihilation operators obey the anticommutation rules, i.e.

$$\{c_{i\sigma}, c_{j\tau}^\dagger\} = c_{i\sigma} c_{j\tau}^\dagger + c_{j\tau}^\dagger c_{i\sigma} = \delta_{i,j} \delta_{\sigma,\tau} \quad (1.22)$$

and

$$\{c_{i\sigma}^\dagger, c_{j\tau}^\dagger\} = \{c_{i\sigma}, c_{j\tau}\} = 0. \quad (1.23)$$

These anticommutation rules provide the properties needed for fermions by construction. First, we notice the Pauli exclusion principle [59, 60] as two fermions cannot occupy the same spin-orbital due to

$$\{c_{i\sigma}^\dagger, c_{i\sigma}^\dagger\} = 2c_{i\sigma}^\dagger c_{i\sigma}^\dagger = 0. \quad (1.24)$$

Furthermore, the anti-symmetry under exchange of two fermions is fulfilled:

$$|\phi_{i\sigma} \phi_{j\tau}\rangle = c_{i\sigma}^\dagger c_{j\tau}^\dagger |-\rangle = -c_{j\tau}^\dagger c_{i\sigma}^\dagger |-\rangle = -|\phi_{j\tau} \phi_{i\sigma}\rangle. \quad (1.25)$$

As a result the construction of any Slater determinant [61] is straightforwardly obtained by using a product of creation operators, i.e.

$$|n_{1\uparrow}n_{1\downarrow} \dots n_{L\uparrow}n_{L\downarrow}\rangle = (c_{1\uparrow}^\dagger)^{n_{1\uparrow}} (c_{1\downarrow}^\dagger)^{n_{1\downarrow}} \dots (c_{L\uparrow}^\dagger)^{n_{L\uparrow}} (c_{L\downarrow}^\dagger)^{n_{L\downarrow}} |-\rangle. \quad (1.26)$$

Second quantization is not unique to fermions; bosonic states can also be represented by a succession of creation operators. For bosons, the creation and annihilation operators b_α^\dagger and b_α respect the commutation rules instead, i.e.

$$[b_\alpha, b_\beta^\dagger] = b_\alpha b_\beta^\dagger - b_\beta^\dagger b_\alpha = \delta_{\alpha,\beta} \quad (1.27)$$

and

$$[b_\alpha^\dagger, b_\beta^\dagger] = [b_\alpha, b_\beta] = 0. \quad (1.28)$$

Although the majority of this dissertation treats electrons which are fermions, we will encounter bosonic states in Chapter 4 where electron pairs are studied as composite bosons with zero spin. For more information on the second quantization, we refer the reader to Ref. [62].

1.2.3. Symmetries

The chemical Hamiltonian in Eq. (1.18) exhibits several symmetries. First, one can easily check that the Hamiltonian is U(1)-symmetric, i.e. the Hamiltonian is invariant under these transformations

$$c^\dagger \rightarrow e^{i\phi} c^\dagger \quad (1.29)$$

$$c \rightarrow e^{-i\phi} c, \quad (1.30)$$

where ϕ is a global phase equal for all creation and annihilation operators. This symmetry is due to the pairwise occurrence of creation and annihilation operators; it gives rise to a conservation of the number of electrons.

Second, the Hamiltonian is also SU(2)-symmetric. The Hamiltonian is invariant when transforming the local spin basis of the electrons as follows

$$c'_\sigma = \sum_{\sigma' \in \{\uparrow, \downarrow\}} U_{\sigma\sigma'} c_{\sigma'} \quad (1.31)$$

$$c'^\dagger_\sigma = \sum_{\sigma' \in \{\uparrow, \downarrow\}} U_{\sigma\sigma'}^* c_{\sigma'}^\dagger \quad (1.32)$$

where $U \in \text{SU}(2)$, i.e.

$$UU^\dagger = U^\dagger U = \mathbb{1} \quad (1.33)$$

and

$$\det(U) = 1. \quad (1.34)$$

The shrewd reader will notice that restricting to $\det(U) = 1$ is rather superfluous; any $U \in U(2)$ would suffice. However, the $U(2)$ -symmetry emerges from the $U(1) \otimes SU(2)$ -symmetry of the Hamiltonian and the simultaneous particle and total spin conservation give rise to the $U(2)$ -symmetry. One can easily check that it is not $U(2)$ but $SU(2)$ that governs spin-symmetries by considering particle-breaking but spin-conserving terms such as $c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger + c_{i\downarrow} c_{i\uparrow}$.

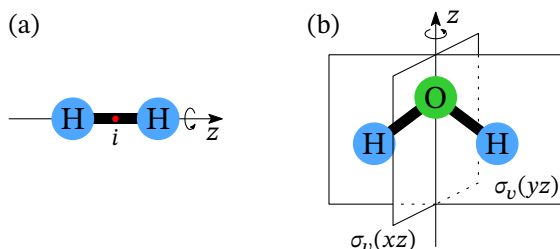


Figure 1.2: The geometry of the H_2 and H_2O molecules. (a) the hydrogen dimer has an inversion center i and rotation-axis z . (b) a water molecule has a rotation axis z for rotations over 180° and two reflection planes $\sigma_v(xz)$ and $\sigma_v(yz)$.

The last symmetry is less obvious when looking at Eq. (1.18) but quite apparent when looking at the system at hand. Molecules can be invariant under rotations, reflections and inversions and the given Hamiltonian built from kinetic terms and Coulomb interactions is invariant under the same set of operations. The exact operations under which the geometry is invariant give rise to the point group symmetry P . For example, H_2 given in Fig. 1.2 is clearly invariant under any rotation along the z -axis and has an inversion center i ; its point group symmetry is $D_{\infty h}$. Another example is H_2O which has a C_{2v} symmetry; it is invariant for rotations over 180° along the z -axis and for reflections along the xz and yz plane. It is obvious that not all molecules have nontrivial point group symmetries, but when they do the eigenstates of the Hamiltonian will transform according to the irreducible representations (irreps) of the point group P .

The most straightforward way to express the point group symmetry in the Hamiltonian is when the spatial orbitals themselves transform according to its irreps. The orbitals involved in the non-zero terms of $V_{ijkl} c_{i\sigma}^\dagger c_{j\tau}^\dagger c_{l\tau} c_{k\sigma}$ need to transform according to the irreps I_i, I_j, I_k and I_l such that $I_i \otimes I_j = I_k \otimes I_l$. This ensures a trivial transformation of the two-electron interactions in the Hamiltonian. One could indeed check that the integrals in Eq. (1.17) are zero

when this is not the case. Similarly, the one-electron integrals T_{ij} will evaluate to zero if $I_i \neq I_j$.

Using symmetries of the system for a better understanding is a much used technique in both quantum physics and chemistry [63–67]. In Chapter 3 we will discuss how the $U(1) \otimes SU(2) \otimes P$ symmetry of the chemical Hamiltonian can be used to push the limits of tensor network methods. Concerning the point group symmetries, we will restrict ourselves to only the simplest ones, i.e. the real abelian point group symmetries C_1 , C_i , C_2 , C_s , D_2 , C_{2v} , C_{2h} and D_{2h} .

1.3. Approximate methods

In the previous section we discussed the chemical Hamiltonian and its appropriate approximations. With the Hamiltonian in Eq. (1.18) one could solve the stationary Schrödinger equation given in Eq. (1.2). However, due to the previously discussed exponential growth of the total Hilbert space, exactly solving the Schrödinger equation becomes intractable when the number of electrons $N = N_\uparrow + N_\downarrow$ or the number of (spatial) orbitals L become too large. Taking into account only the particle conservation of the spin-up and spin-down electrons, the dimension of the total Hilbert space scales as

$$\dim(H) = \binom{L}{N_\uparrow} \binom{L}{N_\downarrow}. \quad (1.35)$$

It is quite clear that further approximations need to be made. In this section, we will give a very brief overview of some much used methods in quantum chemistry.

We can discern between two large classes of methods: the wave function methods and methods relying on the density functional theory (DFT) [68, 69]. While the first method attempts to find the complete wave function of a given chemical system, DFT only considers the electron density. In theory DFT provides an exact study of quantum mechanical systems at a mean-field cost. Unfortunately, a lack of knowledge of the exact exchange-correlation functional, which is needed for the evaluation of the energy, waters DFT down to an approximate method. Many approximations for the exact functional exist, each with their own strengths and weaknesses. A major drawback of DFT is that the needed functional is not systematically improvable to the exact functional [70]; this is in stark contrast with wave function methods.

When discussing the class of wave function methods, one cannot avoid the Hartree-Fock (HF) method [34, 51–53, 71, 72]. This method searches the variationally most optimal Slater determinant within the given basis set. In this

mean-field theory, electrons do not interact instantaneously, instead each electron is only subjected to a mean-field Coulomb interaction with the other electrons. The HF method returns a set of orthonormal occupied and virtual orbitals spanning the complete basis set. Depending on the restrictions imposed the orbitals can be spatially identical for up and down electrons (restricted Hartree-Fock) or different (unrestricted Hartree-Fock) [73].

Since the HF wave function is given by a single Slater determinant, it only takes correlations due to the Pauli repulsion into account. The discrepancy between the HF energy and the exact energy in the given basis set is called the *correlation energy* [74]. Quite often, we further distinguish between strong (static and nondynamical) and weak (dynamical) correlations [75–78]. Both correlations are, of course, electronic by nature, however their origin is quite different. Strong correlations arise through near-degeneracies of several Slater determinants; the single HF Slater determinant fails to qualitatively describe the exact wave function as there are multiple dominant Slater determinants needed. Dynamic correlations are due to the instantaneous Coulomb repulsion between electrons. Typically, the basis set has to be chosen large enough to capture dynamical correlations such that the electrons can relax adequately. Dynamic correlations are expressed as a vast amount of Slater determinants with small amplitudes in the exact solution.

When a system is dominated by dynamical correlations, we can resort to single-reference methods. These post-HF methods start from the HF wave function which is assumed qualitatively correct for the system. Popular techniques for the incorporation of the dynamical correlations are for example the Møller-Plesset perturbation theory (MP) [79], the configuration interaction (CI) method [61, 80] and the coupled cluster (CC) theory [76, 81–83]. The first method includes dynamical correlations through perturbation up to a given order while the other two techniques consider excitations upon the Hartree-Fock solution. Quite commonly, only single and double excitations are considered resulting in CISD and CCSD, but these methods can be extended to the exact solution by increasing the number of excitations⁵. As such, exactly solving the Schrödinger equation within a given basis set is called the full configuration interaction (FCI) method in quantum chemistry.

These methods are ill-suited to study the correct correlations if the reference function on which to perturb or excite is qualitatively incorrect. Using the previously mentioned methods with the HF wave function as reference will provide poor results when strong correlations are present in the system; instead, one could resort to multireference (MR) methods. One of the most well-

⁵At an exponential increasing cost, of course.

known multireference method is the so-called complete active space (CAS) self-consistent field (SCF) method [84–86]. In this method, a subset of relevant orbitals is chosen as active space while the other orbitals are assumed to be either completely occupied (occupied space) or completely empty (virtual space). One optimizes the wave function exactly for the degrees of freedom provided by the active space; this is the CAS part of the algorithm. Once the optimized wave function is obtained, the selected active space is mixed with the occupied and virtual space through orbital rotations, further improving the result. This is the SCF part of the algorithm. These two steps can be iteratively repeated up to convergence. Another multireference method worth to mention is the density matrix renormalization group (DMRG) [87, 88]. This tensor network method solves the system for a chosen active space, albeit approximately. DMRG will be introduced in the next chapter.

The single-reference methods discussed for capturing dynamical correlations can be built on top of a multireference method. Examples are CASPT2 (CAS with perturbation theory up to second order) [89, 90], DMRG-CASPT2 [91, 92], DMRG-NEVPT2 (DMRG with second-order N-electron valence state perturbation theory) [93], p-DMRG [94], MRCC (multireference coupled cluster) [95–97] and DMRG-tailored coupled cluster [98–100]. These methods are well suited for capturing both strong and weak correlations, unfortunately they are numerically quite intensive.

A much more thorough introduction to wave function methods can be found in Refs. [101–103].

1.4. A brief overview of this dissertation

This dissertation will discuss the approximate treatment of the Schrödinger equation through the three legged tree tensor network state (T3NS) [104, 105], an *ab initio* variational wave function method. Just as all tensor network state methods, it can provide an efficient representation of wave functions with low entanglement and it is particularly suited for molecular systems with strong correlations. The T3NS will be introduced in Chapter 2 together with a brief overview of other tensor network techniques. The chapter is concluded by a first cautious numerical comparison between the T3NS and the density matrix renormalization group (DMRG), the first and most widely used tensor network method [87, 88].

As discussed in Section 1.2.3, the chemical Hamiltonian exhibits multiple symmetries. In Chapter 3 we will discuss how one can exploit the symmetries present. The techniques discussed and developed are further used in Chapter 4 for the study of molecules where only a restricted number of broken electron pairs is allowed. The seniority (the number of unpaired electron) is not

a conserved quantity for the chemical Hamiltonian; however, by making the different tensors in the T3NS invariant for the seniority we can easily target Hilbert subspaces corresponding to a restricted seniority number.

The geometry of the tensor network state and the exact mapping of the orbitals has a profound influence on the exact entanglement structures that can be represented. As such, the influence of having a good ordering and geometry cannot be understated for the accuracy and convergence speed of the given tensor network state. In Chapter 5 we will briefly discuss existing methods for ordering orbitals onto a tensor network state and we will present a method for locally permuting the orbitals in the network.

We conclude this dissertation with Chapter 6. A discussion of the conclusions and a general outlook for future research into tensor network states and quantum chemistry is provided.

2

Tensor network states

*I might have thought that the new ideas were correct
if they had not been so ugly.*

P.A.M. Dirac

We provide a pedagogical introduction to tensor network states. The theory and applications of tensor networks are vast and versatile; as such, we will focus specifically on ground state calculations for quantum chemistry, while being forced to merely touch upon (or even completely leave out) a multitude of other techniques established in the field of tensor networks.

We will start by providing a short historical overview of how the most widely used tensor network, the density matrix renormalization group, came into existence. Thereafter, we study the DMRG algorithm by viewing it as a variational optimization within the matrix product state (MPS) manifold.

Next, the main subject of this dissertation is introduced: the three-legged tree tensor network state (T3NS). The T3NS is more versatile than DMRG for encoding entanglement while still being computationally tractable for quantum chemistry. We end with a first numeric comparison of DMRG and T3NS.

Parts of this chapter have been published in Journal of Chemical Theory and Computation **14**, 2026-2033 (2018) [104].

2.1. A concise history

It may seem an odd choice to introduce the main subject of my dissertation with the epigraph by Paul A.M. Dirac given on the previous page; however, all historical overviews need a starting point and I have chosen this as mine. It was reportedly Dirac's response when Freeman Dyson inquired about his opinion on *renormalization* in quantum electrodynamics (QED). The festering problem of divergencies within QED was finally solved by Tomonaga [106], Schwinger [107, 108], Feynman [109–111] and Dyson [112, 113], and the latter probably hoped on some congratulatory words from his former supervisor. Alas, renormalization may have made QED one of the most accurate theories in physics so far¹, but Dirac loathed the *ad hoc* fashion in which divergences canceled out.

It took until the 1970s when Kenneth Wilson removed the *ad hoc*-ness of renormalization and formulated the *renormalization group* [116–119], a framework for the study of physical systems on different scales. Wilson concentrated on problems which have a huge amount of degrees of freedom within regions the size of the correlation length of said problem. The binding of large molecules, the Kondo problem (the scattering of conducting electrons due to a magnetic impurity in a metal), critical phenomena and QED are all such problems. He compared the idea of the renormalization group with hydrodynamics [120]; while the fluid is microscopically characterized by an enormous amount of degrees of freedom, fluid mechanics only needs a few macroscopic variables such as the density $\rho(x)$. The hydrodynamic equations in their turn only depend on these macroscopic variables. To understand the dynamics of the fluid only a coarse resolution of variables such as $\rho(x)$ are needed; luckily, we do not need to know the exact position and orientation of every molecule in the fluid. This is similar in idea to the renormalization group; the original degrees of freedom of the problem should be reduced to a smaller set of *effective* degrees of freedom subject to *effective* interactions.

Let us clarify the renormalization group some more by sketching how Wilson used it to solve the Kondo problem [119]. First, the interaction between the magnetic impurity and the conduction band was approximated as a one-dimensional half-infinite lattice. Wilson tackled this lattice by solving a small sublattice and identifying the relevant effective degrees of freedom of this sublattice. In a next step the Hamiltonian of the sublattice is projected onto these effective degrees of freedom and the sublattice is extended. By repeating this procedure and systematically extending the lattice, Wilson could study the complete half-infinite lattice while only keeping track of a small set of effec-

¹Current predictions of the electron magnetic moment g correspond with experiment up to an astonishing ten significant numbers [114, 115].

tive degrees of freedom. As the selection procedure for the relevant effective degrees of freedom he opted for the lowest-lying eigenstates of the sublattice. This procedure is schematically represented in Fig. 2.1.

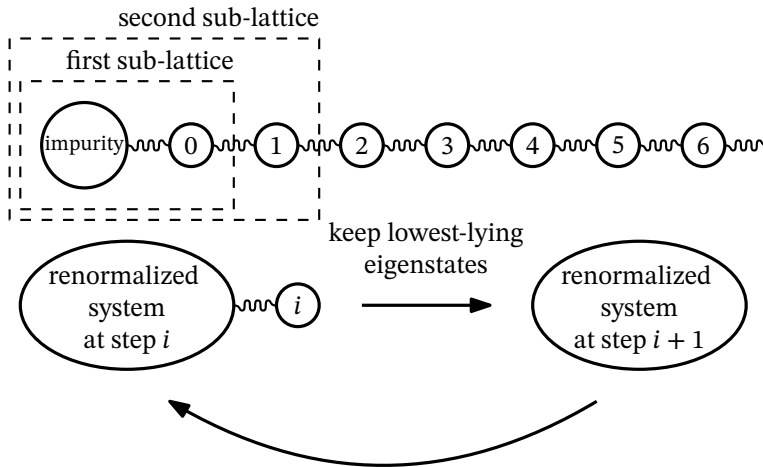


Figure 2.1: Schematic representation of Wilson's renormalization group for the Kondo problem.

Wilson's success in treating the Kondo problem caused quite a stir in the physics community and people hopefully applied the same renormalization group on quantum lattice problems, but to no avail. White and Noack identified that for such problems the low-lying eigenstates of small sublattices aren't important degrees of freedom for the larger whole [121]. It was not the renormalization group but the way effective degrees of freedom were selected for these problems that fell short.

Later on, Steve White successfully identified the right degrees of freedom for these problems as the dominant eigenstates of the reduced density matrix of the ground state [87, 88]; the density matrix renormalization group (DMRG) was born. One can indeed realize that these are the useful degrees of freedom (or renormalized states) by assuming we have the exact ground state of the lattice available and we split the system into two parts. If we now want to reconstruct the ground state with two sets of effective degrees of freedom (one set living on each part) then the most efficient sets are given by the Schmidt decomposition, as seen in Section 1.1.1; this also corresponds with the eigenstates of the corresponding reduced density matrix.

The density matrix renormalization group method proved its usefulness in the simulation of strongly correlated quantum systems, both in condensed matter physics and theoretical chemistry. While initially applied on systems with local Hamiltonians, it didn't take long before it was applied successfully

on systems with long-range interactions, like in momentum space representation (k-DMRG) [122] and quantum chemistry (QC-DMRG) [123].

2

2.2. Modern DMRG

A while after the first formulation of DMRG by White, it became apparent that it corresponds with a variational optimization within the matrix product state (MPS) manifold [124–126]. The better understanding into the machinery of the DMRG algorithm provided a fertile soil for novel algorithms such as the variational unitary matrix product state (VUMPS) [127] and the time-dependent variational principle (TDVP) [128]. It also paved the way for advancements based on extensions of the MPS, the so-called tensor network states. More information on these extensions is provided in Section 2.4.

Although DMRG already proved its power from the start, a true understanding of its success came when Hastings formulated the area law for ground states of local gapped Hamiltonians in one dimension [20]. The area law states that the entanglement between sub-system and environment of such ground states scales with the area of the boundary between the two, and not with the overall volume of the sub-system. When picking a segment out of a one-dimensional system, the entanglement with the rest of the system will not scale with the length of the chosen segment. Specifically for DMRG, this means that the needed number of renormalized states for such systems does not scale with the system size.

In the next sections, we will formally introduce the DMRG algorithm to the reader in the modern MPS language as it is more potent and flexible than White's original formulation.

2.2.1. The matrix product state

Although the focus in this thesis is on the application of tensor network states in a quantum chemical setting, we will introduce the matrix product state (MPS) by using a one-dimensional spin-1/2 chain. The parallel between the MPS and system structure is a welcome plus for a first introduction.

Let us first specify the spin chain at hand; n spin degrees of freedom are placed on a line, and each has a total spin $S = 1/2$. In the S^z -basis, the local Hilbert space of each spin is spanned by the states $\{|\uparrow\rangle, |\downarrow\rangle\}$. The exact wave function can now be written as

$$|\Psi\rangle = \sum_{\{s_i\}} C_{s_1 s_2 s_3 \dots s_n} |s_1 s_2 s_3 \dots s_n\rangle, \quad (2.1)$$

where s_i is one of the local states for spin i , i.e. $s_i \in \{\uparrow, \downarrow\}$ and the complete wave function is given by a linear combination of all possible spin configurations.

This wave function is clearly plagued by the exponential wall of quantum mechanics; the total Hilbert space and configuration interaction (CI) tensor C both grow exponentially with system size. Instead of taming the beast that lies within Eq. (2.1), let us instead try to split it up in digestible chunks. For this, the singular value decomposition can be used; reshaping the CI tensor to a matrix, we can write

$$|\Psi\rangle = \sum_{\{s_i\}} C_{s_1; s_2 s_3 \dots s_n} |s_1 s_2 s_3 \dots s_n\rangle \quad (2.2)$$

$$= \sum_{\{s_i\}} \sum_{\alpha_1} A_{s_1; \alpha_1}^1 \sigma_{\alpha_1} B_{\alpha_1; s_2 s_3 \dots s_n} |s_1 s_2 s_3 \dots s_n\rangle \quad (2.3)$$

$$= \sum_{\{s_i\}} \sum_{\alpha_1} A_{s_1 \alpha_1}^1 C'_{\alpha_1 s_2 s_3 \dots s_n} |s_1 s_2 s_3 \dots s_n\rangle, \quad (2.4)$$

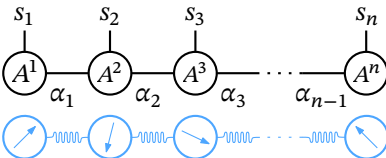
In the last line, we absorbed the singular values into the matrix B forming C' . We can split of the next spin from the newly obtained tensor C' by reiterating this procedure; in fact, by repeating the singular value decomposition, we eventually obtain

$$|\Psi\rangle = \sum_{\{s_i\}} \sum_{\{\alpha_i\}} A_{s_1 \alpha_1}^1 A_{\alpha_1 s_2 \alpha_2}^2 A_{\alpha_2 s_3 \alpha_3}^3 \dots A_{\alpha_{n-1} s_n}^n |s_1 s_2 s_3 \dots s_n\rangle. \quad (2.5)$$

Each local degree of freedom (spin) of the system has an associated tensor and each state of this local degree of freedom has an associated matrix. For example the matrix $A_{\alpha_1 \uparrow \alpha_2}^2$ is associated with the second spin in an up state. To obtain the coefficient corresponding with a certain spin configuration we can take the product of the appropriate matrices; *we have obtained the matrix product state ansatz*.

Decomposing the CI tensor into this MPS has introduced the set $\{\alpha_i\}$ as new indices. These indices are called the *virtual bonds* or *virtual degrees of freedom* as opposed to the physical degrees of freedom $\{s_i\}$. They take care of the interconnection between the different spins; without them, a simple product state would have been obtained without any correlations between the spins.

We can graphically depict the MPS in Eq. (2.5) as follows:



$$\quad (2.6)$$

Under the MPS, we included a representation of the Heisenberg system we are studying; it further accentuates the correspondence between the structure

of the MPS and of the system. Every MPS tensor is represented by a circle while the indices of the tensors are represented by edges. The virtual bonds interconnect the different tensors while the physical degrees of freedom (or bonds) stay loose. Bonds which connect two tensors represent a contraction of the tensors over said index. We will heavily rely on graphical representations of tensor networks and contractions as it improves a lot on the readability of the needed algebraic equations.

Although the beast is divided in chunks, we did not yet tame it. When closely examining the dimensionality of the newly introduced virtual bond indices, its exponential growth towards the middle of the tensor train² is apparent; we did not avoid the exponential wall, we merely moved it. A way out of this predicament does present itself; we have to truncate the SVD. By keeping the largest singular values at each SVD step, a controlled approximation of the wave function can be made. The dimension of virtual bonds kept after truncation is called, quite unsurprisingly, the *virtual bond dimension*. In Eq. (2.5), the virtual bond dimension is given by the dimensionality of the indices α_i ; it is a central concept in tensor network methods as it is the primary parameter influencing the accuracy of the method.

The gauge freedom

Starting from the exact wave function given in Eq. (2.1), we have derived the MPS in Eq. (2.5). However, the MPS for a given wave function is not unique. We can easily change the individual tensors without changing the complete wave function itself as follows:

$$\begin{array}{c} | \\ \circlearrowleft A \\ | \end{array} \text{---} \begin{array}{c} | \\ \circlearrowleft B \\ | \end{array} = \begin{array}{c} | \\ \circlearrowleft A \\ | \end{array} \text{---} \begin{array}{c} | \\ \circlearrowleft X \\ | \end{array} \text{---} \begin{array}{c} | \\ \circlearrowleft X^{-1} \\ | \end{array} \text{---} \begin{array}{c} | \\ \circlearrowleft B \\ | \end{array} = \begin{array}{c} | \\ \circlearrowleft A' \\ | \end{array} \text{---} \begin{array}{c} | \\ \circlearrowleft B' \\ | \end{array} \quad (2.7)$$

where X is any invertible matrix. This gauge freedom living on the virtual bonds can be used to bring the MPS in the so-called *canonical form*.

The canonical form

By construction through the SVD, all tensors except the last tensor A^n satisfy

$$\sum_{s_i \alpha_{i-1}} A_{\alpha_{i-1} s_i \alpha_i}^i \bar{A}_{\alpha_{i-1} s_i \alpha_i}^i = \delta_{\alpha_i \alpha_i'} \quad (2.8)$$

The MPS in Eq. (2.5) is not just any MPS, it is one written in the so-called *canonical form*. We call an MPS canonical if there is one tensor defined as *orthogonality center*, while all the other tensors are orthogonal with respect to contraction over all indices except the virtual degree of freedom leading to the

²It may come as no surprise that this is another name for the MPS [129].

orthogonality center. Formulating it in another way: if c is chosen as orthogonality center, the other tensors should fulfill

$$\begin{array}{c} \bar{A}^i \\ \alpha_{i-1} \quad s_i \\ A^i \\ \alpha_i \end{array} = \begin{array}{c} \alpha'_i \\ \alpha_i \end{array} \quad \forall i < c \quad (2.9)$$

$$\begin{array}{c} \bar{A}^i \\ \alpha'_{i-1} \quad s_i \\ A^i \\ \alpha_{i-1} \end{array} = \begin{array}{c} \alpha'_{i-1} \\ \alpha_{i-1} \end{array} \quad \forall i > c \quad (2.10)$$

where the Kronecker delta of Eq. (2.8) is represented as a continuous line.

The tensor assigned as orthogonality center can be changed by exploiting the gauge freedom of the MPS. We can QR-decompose the current orthogonality center and absorb the resulting triangular R-matrix into a neighboring tensor, making this tensor the new orthogonality center, i.e.

$$\begin{array}{c} s_i \quad s_{i+1} \\ \dots \quad \bar{A}^i \quad \alpha_i \quad \bar{A}^{i+1} \quad \alpha_{i+1} \quad \dots \\ \dots \quad \bar{A}^i \quad \alpha'_{i-1} \quad R \quad \alpha_i \quad \bar{A}^{i+1} \quad \alpha_{i+1} \quad \dots \\ \dots \quad \bar{A}^i \quad \alpha'_{i-1} \quad \bar{A}^{i+1} \quad \alpha_{i+1} \quad \dots \end{array} \quad (2.11)$$

We colored the current orthogonality center at each step, while the uncolored tensors are suitably orthogonalized.

The canonical form provides some notable advantages, as is often the case for canonical forms in physics and mathematics. First, it increases the numerical stability of the optimization and evaluation of properties. As an illustration, let us evaluate the expectation value of S_n^z for the MPS:

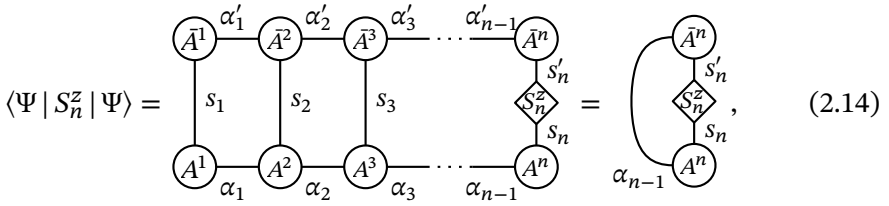
$$\begin{aligned} \langle \Psi | S_n^z | \Psi \rangle &= \sum_{\{s_i\} \{ \alpha_i \} \{ \alpha'_i \} s'_n} A^1_{s_1 \alpha_1} \bar{A}^1_{s_1 \alpha'_1} A^2_{\alpha_1 s_2 \alpha_2} \bar{A}^2_{\alpha'_1 s_2 \alpha'_2} A^3_{\alpha_2 s_3 \alpha_3} \bar{A}^3_{\alpha'_2 s_3 \alpha'_3} \dots \\ &\quad \delta_{s_n, s'_n} (-1)^{\delta_{s_n, \downarrow}} A^n_{\alpha_{n-1} s_n} \bar{A}^n_{\alpha'_{n-1} s'_n}, \end{aligned} \quad (2.12)$$

as all tensors A^i but the last are orthogonal with respect to contraction over s_i and α_{i-1} , we have

$$= \sum_{s_n s'_n \alpha_{n-1}} \delta_{s_n, s'_n} (-1)^{\delta_{s_n, \downarrow}} A^n_{\alpha_{n-1} s_n} \bar{A}^n_{\alpha_{n-1} s'_n}. \quad (2.13)$$

Graphically this is depicted as

2



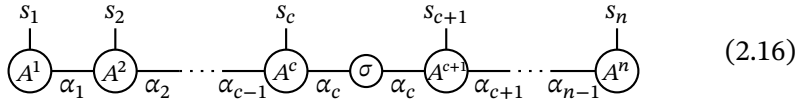
(2.14)

where we introduced the S_n^z operator graphically. The simplification and additional numerical stability due to the canonical form is quite apparent; instead of numerous consecutive tensor contractions over the different A^i and their adjoints, only the last tensors are involved in the evaluation. We just studied the special case where an operator only acts upon the local Hilbert space of one tensor which we made the current orthogonality center, but similar, albeit less pronounced, simplifications also occur for the evaluation of other few-body operators.

Another advantage of the canonical form is the readiness for the extraction of Schmidt values. As discussed in Section 1.1.1, these provide a measurement of the entanglement between a part of the system and its environment. Now, let us assume the orthogonality center is at spin c . An appropriate singular value decomposition

$$A_{\alpha_{c-1}s_c\alpha_c}^c = A_{\alpha_{c-1}s_c\alpha'_c}^c \sigma_{\alpha'_c} V_{\alpha'_c\alpha_c} \quad (2.15)$$

followed by a contraction of $V_{\alpha'_c\alpha_c}$ with A^{c+1} results in the following wave function:



where σ is a diagonal matrix with the singular values as elements and

$$\begin{array}{c} \textcircled{\bar{A}^i} - \alpha'_i \\ \alpha_{i-1} \text{---} | \text{---} s_i \\ \textcircled{A^i} - \alpha_i \end{array} = \begin{array}{c} \alpha'_i \\ \alpha_i \end{array} \quad \forall i \leq c \quad (2.17)$$

$$\begin{array}{c} \alpha'_{i-1} \text{---} \textcircled{\bar{A}^i} \\ | \text{---} s_i \\ \alpha_i \text{---} \textcircled{A^i} \\ \alpha_{i-1} \text{---} \end{array} = \begin{array}{c} \alpha'_{i-1} \\ \alpha_{i-1} \end{array} \quad \forall i > c \quad (2.18)$$

i.e. all the tensors are properly orthogonalized. Investigating the wave function in Eq. (2.16), the Schmidt decomposition between the first c spins and the last

spins is retrieved; it is given by

$$|\Psi\rangle = \sum_{\alpha_c} \sigma_{\alpha_c} |\varphi_{\alpha_c}\rangle |\chi_{\alpha_c}\rangle \quad (2.19)$$

with

$$|\varphi_{\alpha_c}\rangle = \sum_{\{s_i\}} \begin{array}{c} s_1 \\ | \\ A^1 \\ | \\ \alpha_1 \end{array} \text{---} \begin{array}{c} s_2 \\ | \\ A^2 \\ | \\ \alpha_2 \end{array} \text{---} \dots \text{---} \begin{array}{c} s_c \\ | \\ A^c \\ | \\ \alpha_c \end{array} \quad |s_1 s_2 \dots s_c\rangle \quad (2.20)$$

$$|\chi_{\alpha_c}\rangle = \sum_{\{s_i\}} \begin{array}{c} s_{c+1} \\ | \\ A^{c+1} \\ | \\ \alpha_{c+1} \end{array} \text{---} \dots \text{---} \begin{array}{c} s_{n-1} \\ | \\ A^{n-1} \\ | \\ \alpha_{n-1} \end{array} \text{---} \begin{array}{c} s_n \\ | \\ A^n \\ | \\ \alpha_n \end{array} \quad |s_{c+1} \dots s_{n-1} s_n\rangle. \quad (2.21)$$

Due to the orthogonality of the different tensors, it is easy to check that both $\{|\varphi_{\alpha_c}\rangle\}$ and $\{|\chi_{\alpha_c}\rangle\}$ are orthogonal sets. As mentioned in Section 1.1.1, the Schmidt states $|\varphi_{\alpha_c}\rangle$ and $|\chi_{\alpha_c}\rangle$ correspond with eigenstates of the appropriate reduced density matrix while the Schmidt values are the square root of the corresponding eigenvalues; these Schmidt states are thus the renormalized states of the original DMRG formulation. With some liberty of words, we will talk in the future about the renormalized states ‘living’ in the virtual bonds.

The Schmidt decomposition of other Hilbert space partitions is obtained with ease as long as the partitioning is natural to the MPS, i.e. when looking at the linear MPS in Eq. (2.6) the partition of the Hilbert space corresponds with snapping the MPS in two pieces. If the partitioning does not correspond with this ‘snapping’, Schmidt values can still be obtained, albeit not as straightforward.

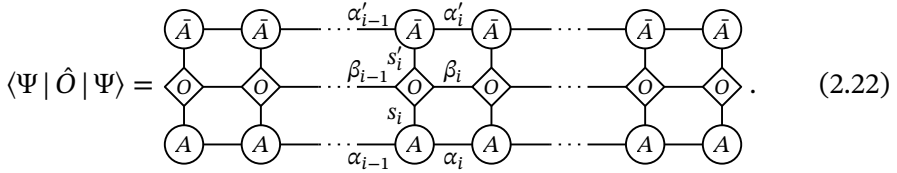
The maximal entanglement that can be encoded by the MPS clearly depends on the virtual bond dimensions. Looking back at one-dimensional systems following the area law, we do not expect a scaling of the needed bond dimensions with increasing system size for such systems. For a given accuracy, we have essentially a polynomial scaling algorithm as long as the needed bond dimensions do not scale exponentially with system size.

2.2.2. Matrix product operators

Let us now consider how the evaluation of $\langle\Psi|\hat{O}|\Psi\rangle$ for any given operator \hat{O} can happen naturally when the bra and ket are written as an MPS. In the spirit of the MPS, we could attempt to write the operator itself as a tensor network.

Making the idea more concrete, our aim is to obtain something like this:

2



Here, the upper and lower layer of tensors represent the bra and ket wave function in MPS form respectively. The middle layer is a so-called matrix product operator (MPO) representation of \hat{O} ; it represents the operator as a linear network of individual tensors. In this equation, we have left out most of the index labels as to not clutter the equation. Besides the two physical indexes to connect to the bra and ket MPS wave function, the MPO tensors also have virtual MPO indices (β) to interconnect the different MPO tensors. By contracting the tensor network in Eq. (2.22) we can evaluate the expectation value.

Just as the MPS is particularly suited for wave functions respecting the one-dimensional area law, some operators are easier and more efficiently written in an MPO than others. As an example we will discuss the one-dimensional Heisenberg Hamiltonian:

$$\hat{H} = J \sum_{i=1}^{n-1} \hat{S}_i \cdot \hat{S}_{i+1} = \sum_{i=1}^{n-1} \frac{J}{2} \hat{S}_i^+ \hat{S}_{i+1}^- + \frac{J}{2} \hat{S}_i^- \hat{S}_{i+1}^+ + J \hat{S}_i^z \hat{S}_{i+1}^z. \quad (2.23)$$

Its short range and its linear geometry make it particularly suited for transforming it into an MPO. Each term in this Hamiltonian is essentially an MPO with a virtual bond of one, e.g.

$$\underbrace{\hat{1} \otimes \hat{1} \otimes \dots \otimes \hat{1}}_{i-1 \text{ times}} \otimes \frac{J}{2} \hat{S}_i^+ \otimes \hat{S}_{i+1}^- \otimes \underbrace{\hat{1} \otimes \dots \otimes \hat{1}}_{n-i-1 \text{ times}}.$$

A summation of all these terms would not provide a compact MPO representation of the complete Hamiltonian in Eq. (2.23); we need to tackle this problem more creatively. Construction of the needed MPO through finite-state machines [130–132] results in the most compact formulation of the Hamiltonian

possible; namely all bulk MPOs are given by

$$\hat{O}_{\beta_{i-1}\beta_i}^i = \begin{bmatrix} \hat{1} & \hat{S}_i^+ & \hat{S}_i^- & \hat{S}_i^z & 0 \\ 0 & 0 & 0 & 0 & \frac{J}{2}\hat{S}_i^- \\ 0 & 0 & 0 & 0 & \frac{J}{2}\hat{S}_i^+ \\ 0 & 0 & 0 & 0 & J\hat{S}_i^z \\ 0 & 0 & 0 & 0 & \hat{1} \end{bmatrix}$$

and the first and last MPO are given by

$$\hat{O}_{\beta_1}^1 = \begin{bmatrix} \hat{1} & \hat{S}_1^+ & \hat{S}_1^- & \hat{S}_1^z & 0 \end{bmatrix} \quad \hat{O}_{\beta_{n-1}}^n = \begin{bmatrix} 0 \\ \frac{J}{2}\hat{S}_n^- \\ \frac{J}{2}\hat{S}_n^+ \\ J\hat{S}_n^z \\ \hat{1} \end{bmatrix}.$$

The virtual bond indexes β_i can be seen as indexes for the different states of the finite state machines, i.e. $\beta_i \in \{\mathbb{1}, +, -, z, H\}$. The state $\mathbb{1}$ means only $\hat{1}$ operators have been acting on spins to the left of this bond, the state H collects all completed Hamiltonian terms and the states $\{+, -, z\}$ represent how the corresponding \hat{S} operator acted on the previous spin. The physical bond indices are not explicitly shown in the above MPOs; they are part of the operators composing the MPOs. It is easy to check that contracting the given MPOs results in the full Hamiltonian in Eq. (2.23). Although this MPO construction looks quite straightforward, it is worth pointing out that it is not always the case; other, more complicated operators such as the quantum chemical Hamiltonian can also be efficiently written as an MPO, albeit with more elaborate techniques [133–135].

The representation of the Hamiltonian is far from the only usage of the MPO; in fact the MPO was first introduced as a means for calculating finite-temperature density matrices [136]. For the reader who wants to discover the many faces of the MPO, we refer to Ref. [137].

2.2.3. The DMRG algorithm

During the DMRG optimization, the tensors of the MPS are usually optimized one or two at a time. We will start with the discussion of the two-site optimization as the one-site optimization is quite similar.

The MPS can be optimized by minimizing the following Lagrangian:

$$\mathcal{L} = \langle \Psi_{\text{MPS}} | \hat{H} | \Psi_{\text{MPS}} \rangle - \lambda \langle \Psi_{\text{MPS}} | \Psi_{\text{MPS}} \rangle, \quad (2.24)$$

This Lagrangian can also be graphically depicted as

$$\mathcal{L} = \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \text{---} \diamond \text{---} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array} - \lambda \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array} \quad (2.25)$$

where we have dropped the labels of the bonds and the tensors. The upper and lower circular layer represent the bra and ket MPS respectively. The square layer represents again the MPO.

Now, imagine we want to minimize this Lagrangian with respect to tensor i and $i + 1$. Let us first contract the tensors A^i and A^{i+1} to a new, larger, tensor B and instead minimize the Lagrangian with respect to the elements of this tensor. The Lagrangian to optimize is now given by

$$\mathcal{L} = \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \text{---} \bar{B} \text{---} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \text{---} B \text{---} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array} - \lambda \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \text{---} \bar{B} \text{---} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \text{---} B \text{---} \text{---} \circ \text{---} \end{array} \quad (2.26)$$

$$= \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \bar{B} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} B \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array} - \lambda \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \bar{B} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \circ \text{---} B \text{---} \circ \text{---} \end{array} \quad (2.27)$$

where we assumed that B is the current orthogonality center and the other tensors are suitably orthogonalized.³ Further, we introduced a new tensor representing a partial contraction of the bra and ket MPS with the MPO, i.e.

$$\underbrace{\begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array}}_{i-1 \text{ left orthogonal MPSs}} = \left(\text{---} \right), \quad \underbrace{\begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array}}_{n-i-1 \text{ right orthogonal MPSs}} = \left. \text{---} \right). \quad (2.28)$$

Minimization of this Lagrangian with respect to \bar{B} is now given by

$$\frac{\partial \mathcal{L}}{\partial \bar{B}} = \begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \bar{B} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} B \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array} - \lambda \left(\text{---} \bar{B} \text{---} \right) = 0. \quad (2.29)$$

Due to the canonical form of the MPS, an ordinary eigenvalue problem is obtained, i.e.

$$\begin{array}{c} \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} \bar{B} \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \\ \text{---} \diamond \text{---} B \text{---} \diamond \text{---} \\ \text{---} \circ \text{---} \text{---} \circ \text{---} \end{array} = \lambda \left(\text{---} \bar{B} \text{---} \right) \quad (2.30)$$

³We can always use the gauge freedom to make B the orthogonality center.

Or algebraically

$$H_{\text{eff}} \vec{B} = \lambda \vec{B}, \quad (2.31)$$

where

$$H_{\text{eff}} = \left(\text{---} \text{---} \text{---} \right). \quad (2.32)$$

In Eq. (2.31), we flattened (or vectorized) the tensor elements B . In Eq. (2.32) The effective Hamiltonian is constructed by the tensors given in Eq. (2.28) and the MPOs belonging to the physical degrees of freedom of B ; it is a partial contraction of the bra and ket MPS with the MPO while excluding the B and \bar{B} tensors. Solving the resulting ordinary eigenvalue problem results in an optimal two-site tensor B' .

However, we do not need the optimal two-site tensor but the optimal tensors A^i and A^{i+1} . The dimension of the virtual bond connecting the two tensors is generally smaller than the maximal allowed one; generally speaking, B' cannot be exactly decomposed in A^i and A^{i+1} with the imposed maximal bond dimension. However, through a truncated singular value decomposition where we discard the smallest singular values we can decompose B' with a minimal error, see also Appendix A. The truncation immediately provides us with a measure of the approximations made. This procedure of approximating B' corresponds exactly with the original formulation of DMRG, i.e. only the most important eigenstates of the reduced density matrix are kept.

Renormalizing the obtained singular values and, for example, contracting them in A^{i+1} leaves the MPS in the canonical form with the orthogonality center at A^{i+1} . The same optimization can be repeated for A^{i+1} and A^{i+2} and so on. Once the rightmost tensor is reached our right sweep is finished and we can start the left sweep. This sweeping procedure is repeated until a predefined convergence criterion is reached.

When creating the effective Hamiltonian, it is good to note that the partial contractions in Eq. (2.28) should not be calculated from scratch every step. One partial contraction can be recycled from an old sweep while the other one is easily updated with the new optimized tensor as follows:

The diagram illustrates the contraction of three tensors: \bar{A}^i , O^i , and A^i . On the left, \bar{A}^i (top circle), O^i (middle diamond), and A^i (bottom circle) are connected by vertical lines. \bar{A}^i and A^i are also connected by a horizontal line on their left side. An arrow points to the right, where the result is a single tensor with three physical indices: α_i' (top), β_i (middle), and α_i (bottom).

$$\left(\begin{array}{c} \bar{A}^i \\ O^i \\ A^i \end{array} \right) \rightarrow \left(\begin{array}{c} \alpha_i' \\ \beta_i \\ \alpha_i \end{array} \right). \quad (2.33)$$

The one-site optimization for DMRG is completely equivalent with the sole exception that the optimal one-site tensor A^i is directly searched instead of an

intermediate two-site tensor. Once obtained, the wave function can be brought in the next suitable canonical form by either a QR decomposition or an LQ decomposition. A subtle difference between the two types of optimization exists; while the one-site optimization finds at each optimization step the most optimal tensor, the two-site optimization finds the two tensors which are closest to the most optimal two-site tensor. This is not necessarily the same result.⁴

One-site optimization is, per sweep, faster than the two-site optimization. However, the latter does have its merit. It is less prone to get stuck in local minima, although there exist modifications of the one-site optimization that also tend to avoid local minima [138–140]. Also, when exploiting symmetries of the system during the DMRG optimization, it allows an automatic reallocation of the renormalized states to the most important irreps of the symmetry. See Chapter 3 and more specifically Section 3.4.5 for more information.

2.3. DMRG in quantum chemistry

The chemical Hamiltonian, with its long range Coulomb interaction and the three-dimensionality of most molecules, does not seem to be an ideal system for the MPS ansatz. Regardless, DMRG has solidified its name in the quantum chemical community having reaped success as a highly accurate solver for systems with substantial static correlations, such as transition metal complexes [91, 141–153]. Having introduced the MPS and DMRG in the previous sections, it is now time to study the quantum chemical variant of the DMRG algorithm.

When defining the MPS, the total Hilbert space needs to be rewritten as a product of local Hilbert spaces. For the one-dimensional Heisenberg system in Section 2.2.1 the partitioning in local physical degrees of freedom was evident; for the quantum chemical Hamiltonian it already provides a first hurdle. As discussed in Section 1.2.1, quantum chemical calculations are normally not performed in a Hilbert space built from completely local basis functions such as a three-dimensional grid; instead orbitallike basis sets are used. When working in these basis sets we lose the notion of locality of the physical degrees of freedom partially or completely. This loss of locality has two major implications.

First, the two-body Coulomb interactions are effectively four-point interactions in a non-local basis. In contrast with the previously studied Heisenberg Hamiltonian in Eq. (2.23) where the number of terms scaled linearly with the number of spins, the chemical Hamiltonian scales *quartic* with the number of used orbitals. The utmost care has to be taken into efficiently evaluating this

⁴See Appendix A for an example where the obtained wave functions are vastly different.

Hamiltonian during optimization, failing to do so can horribly influence the complexity of the algorithm. We will study this more closely in Section 2.3.1.

Second, there is a lot of freedom in which orbitals to use and how to order them on the MPS. A good choice and order can widely influence the convergence speed and accuracy of the DMRG algorithm. Do we work with canonical orbitals, localized orbitals or something else? How do we order these orbitals onto a linear MPS? We will briefly touch on this in Section 2.6 and further investigate it in Chapter 5. One decision will be made here however. We will assign to each MPS tensor a spatial orbital; its local Hilbert space is spanned by $n_i \in \{-, \uparrow, \downarrow, \uparrow\downarrow\}$, i.e. each orbital can be empty, singly occupied with an up- or down-electron or doubly occupied. A partitioning of the total Hilbert space on the level of the spin orbitals is also valid, i.e. every orbital is either empty or singly occupied by an electron with appropriate spin. However, we will solely resort to the spatial orbital partitioning in this dissertation as it eases the usage of the $SU(2)$ -symmetry in Chapter 3. In this case, the MPS is given by:

$$\begin{array}{ccccccc}
 n_1 & n_2 & n_3 & & n_{L-2} & n_{L-1} & n_L \\
 | & | & | & & | & | & | \\
 \circ & \circ & \circ & \dots & \circ & \circ & \circ \\
 | & | & | & & | & | & | \\
 \alpha_1 & \alpha_2 & \alpha_3 & & \alpha_{L-3} & \alpha_{L-2} & \alpha_{L-1}
 \end{array} \quad (2.34)$$

where n_i are the occupations of the L different orbitals.

2.3.1. Renormalized operators

Using MPOs for representing operators when the wave function itself is represented by an MPS looks quite natural, but the connection between DMRG and the MPS was not known from day one. Even now a lot of DMRG implementations, especially in quantum chemistry, don't use MPOs but resort to so-called *renormalized operators*. This way of handling the Hamiltonian efficiently was pioneered in the early years of DMRG [122, 123] and arises organically when formulating DMRG in the original language of renormalized states. In this section, we will start with introducing the concept of renormalized operators for quantum chemistry. Once introduced, we will explain its connection with the MPO formalism and why it still has its merit to this day. Although the MPO formalism is well known and there exists MPO forms of the chemical Hamiltonian, there are still new projects resorting to the old concept of renormalized operators [154, 155]. In fact, the T3NS implementation developed for this dissertation also uses renormalized operators [156].

We will disregard the fermionic signs completely in the discussion to come; this is merely a sketch of the renormalized operator formalism. In Section 2.7, we will discuss how to incorporate the fermionic sign. We will also only focus on the four-point terms.

As discussed in Section 2.2.3, the DMRG algorithm optimizes one or two tensors at a time and only needs the effective Hamiltonian with respect to these tensors; we need to project the Hamiltonian on the Hilbert space spanned by the local physical degrees of freedom of the optimized tensors and the renormalized states living in their virtual bonds. In the MPO framework the complete Hamiltonian was projected onto this subspace through Eq. (2.32). Let us now instead project each term of the Hamiltonian separately. We first define the subspace to project on, this subspace is composed of the left and right renormalized states $|\alpha_{i-1}\rangle$ and $|\alpha_{i+1}\rangle$ and the physical degrees of freedom of the two sites we are currently optimizing $|n_i n_{i+1}\rangle$:

$$(2.35)$$

Looking back at the chemical Hamiltonian in Eq. (1.18) we note that up to four-point terms occur. For the terms in the chemical Hamiltonian it is possible that either zero, one, two, three or four creation or annihilation operators act on the left renormalized states, and similarly for the right renormalized states and the local physical states. For example, let us examine the term $V_{12iL} c_{1\sigma}^\dagger c_{2\tau}^\dagger c_{L\tau} c_{i\sigma}$ and project it on the subspace $|\alpha_{i-1}\rangle \otimes |n_i n_{i+1}\rangle \otimes |\alpha_{i+1}\rangle$; we have to calculate

$$(2.36)$$

and

$$(2.37)$$

Once again, we have neglected the fermionic nature of the wave function and the operators; we delay its discussion to Section 2.7. These terms are called the renormalized operators as they are projections of the operators onto the renormalized states. They suffice for the construction of the projection of $V_{12iL} c_{1\sigma}^\dagger c_{2\tau}^\dagger c_{L\tau} c_{i\sigma}$ on the subspace $|\alpha_{i-1}\rangle \otimes |n_i n_{i+1}\rangle \otimes |\alpha_{i+1}\rangle$.

Complementary renormalized operators

Naively it seems that we need to calculate $\mathcal{O}(i^4)$ left renormalized operators, $\mathcal{O}((L-i)^4)$ right renormalized operators and $\mathcal{O}(L^4)$ different combinations of renormalized operators to form the effective Hamiltonian. Of course we can do better. To do so we need *complementary* renormalized operators. For starters, we can group all completed Hamiltonian terms into one renormalized operator, i.e.

$$\begin{array}{c} \tilde{\alpha}_{i-1} \\ \text{---} \\ \boxed{H} \\ \text{---} \\ \alpha_{i-1} \end{array} = \sum_{\substack{a,b,c,d < i \\ \sigma,\tau}} V_{abcd} \begin{array}{c} \tilde{\alpha}_{i-1} \\ \text{---} \\ \boxed{c_{a\sigma}^\dagger c_{b\tau}^\dagger c_{d\tau} c_{c\sigma}} \\ \text{---} \\ \alpha_{i-1} \end{array} \quad (2.38)$$

$$\begin{array}{c} \tilde{\alpha}_{i+1} \\ \text{---} \\ \boxed{H} \\ \text{---} \\ \alpha_{i+1} \end{array} = \sum_{\substack{a,b,c,d > i+1 \\ \sigma,\tau}} V_{abcd} \begin{array}{c} \tilde{\alpha}_{i+1} \\ \text{---} \\ \boxed{c_{a\sigma}^\dagger c_{b\tau}^\dagger c_{d\tau} c_{c\sigma}} \\ \text{---} \\ \alpha_{i+1} \end{array} \quad (2.39)$$

Similarly we can also group the terms where two or three operators acted on the left or right renormalized states, for example:

$$\begin{array}{c} \tilde{\alpha}_{i-1} \\ \text{---} \\ \boxed{Q_{d\uparrow}} \\ \text{---} \\ \alpha_{i-1} \end{array} = \sum_{a,b,c < i} [V_{abdc} - V_{abcd}] \begin{array}{c} \tilde{\alpha}_{i-1} \\ \text{---} \\ \boxed{c_{a\uparrow}^\dagger c_{b\uparrow}^\dagger c_{c\uparrow}} \\ \text{---} \\ \alpha_{i-1} \end{array} \quad (2.40)$$

$$+ [V_{abdc} + V_{bacd}] \begin{array}{c} \tilde{\alpha}_{i-1} \\ \text{---} \\ \boxed{c_{a\uparrow}^\dagger c_{b\downarrow}^\dagger c_{c\downarrow}} \\ \text{---} \\ \alpha_{i-1} \end{array}$$

$$\begin{array}{c} \tilde{\alpha}_{i+1} \\ \text{---} \\ \boxed{P_{d\uparrow c\uparrow}} \\ \text{---} \\ \alpha_{i+1} \end{array} = \sum_{a,b > i+1} [V_{abcd} - V_{abdc}] \begin{array}{c} \tilde{\alpha}_{i+1} \\ \text{---} \\ \boxed{c_{a\uparrow}^\dagger c_{b\uparrow}^\dagger} \\ \text{---} \\ \alpha_{i+1} \end{array} \quad (2.41)$$

The notation of the complimentary renormalized operator in Eq. (2.40) denotes that this operator should be combined with the operator $c_{d\uparrow}$, with d an orbital not contained in the left renormalized states. Similarly for Eq. (2.41), the operator should be combined with $c_{d\uparrow} c_{c\uparrow}$. We need to construct complimentary renormalized operators such as in Eq. (2.40) for both the left as the

right renormalized bases. Complimentary renormalized operators which are sums over terms with two operators as in Eq. (2.41) only have to be constructed on one set of renormalized states. Normally it is the longest side of the MPS chain that is chosen as to get as much speedup as possible.

By using these complimentary renormalized operators, the complexity of constructing the effective Hamiltonian scales quadratically with the number of orbitals. The number of needed renormalized operators also scales quadratically; we only need to store the single and double operators and the complimentary operators.

Updating renormalized operators

Updating preexisting renormalized operators is imperative for an efficient algorithm. Just as for MPOs (see Eq. (2.33)), it is quite straightforward to do so, e.g.:

$$\begin{array}{c}
 \tilde{\alpha}_{i-1} \quad \tilde{A}^i \quad \tilde{\alpha}_i \\
 \boxed{c_{a\uparrow}^\dagger} \quad \text{---} \quad \text{---} \\
 \alpha_{i-1} \quad A^i \quad \alpha_i
 \end{array}
 =
 \begin{array}{c}
 \tilde{\alpha}_i \\
 \boxed{c_{a\uparrow}^\dagger} \\
 \alpha_i
 \end{array}
 \quad (2.42)$$

For updating the complimentary renormalized operators we also have to take into account the new contributions, e.g.

$$\begin{array}{c}
 \tilde{\alpha}_{i-1} \quad \tilde{A}^i \quad \tilde{\alpha}_i \\
 \boxed{Q_{i\uparrow}} \quad \text{---} \quad \text{---} \\
 \alpha_{i-1} \quad A^i \quad \alpha_i \\
 \text{---} \quad \text{---} \\
 \quad \quad \quad \diamond c_{i\uparrow}
 \end{array}
 \rightarrow
 \begin{array}{c}
 \tilde{\alpha}_i \\
 \boxed{H} \\
 \alpha_i
 \end{array}
 \quad (2.43)$$

The result of contracting $Q_{i\uparrow}$ and $c_{i\uparrow}$ should be added to the complimentary operator collecting the completed Hamiltonian terms.

MPOs versus renormalized operators

The attentive reader may have noticed there are quite some parallels between MPOs and renormalized operators although they stem from different viewpoints of DMRG. In fact, when only an effective Hamiltonian is needed, the complete Hamiltonian in MPO form is never needed; instead a partial contraction of the MPO with the bra and ket MPS suffices, as given in Eq. (2.28).

One could imagine that the MPO virtual bond indices of these contracted MPOs actually label the different renormalized operators. Appending new MPOs to these contracted MPOs then encodes the instructions for updating the renormalized operators and forming the effective Hamiltonian.

So why is there still merit in renormalized operators, although it can be seen as a concealed MPO formulation⁵? When combining quantum chemical DMRG with the MPO formalism, the sparsity of the MPOs has to be explicitly exploited; failing to do so can result in a suboptimal computational complexity. When only the partial traces of the MPOs are needed and not the MPOs explicitly – as is the case for a ground state optimization – it is easier to immediately work with renormalized operators to get the correct scaling. When writing the quantum chemical Hamiltonian in MPO format the MPO bond dimension scales quadratic with L [133, 134]; not surprisingly the same scaling as the number of renormalized operators. Attempting to work with this MPO as a dense tensor will result in a cost of $\mathcal{O}(L^4)$ to construct the effective Hamiltonian given in Eq. (2.32); also updating the partial contracted MPOs will cost $\mathcal{O}(L^4)$ per update. This is far from the lowest cost possible, as we will see in Section 2.3.2. Only by carefully taking the sparsity of the MPOs into account is the optimal scaling obtained [133, 134].

For a thorough discussion of the parallels and differences between MPOs and renormalized operators, we refer the reader to Ref. [134].

2.3.2. Resource costs of the algorithm

Using renormalized operators for the QC-DMRG algorithm, we can make an assessment of the resource costs of the algorithm. The number of orbitals (and sites) is denoted by L while D is the virtual bond dimension of the MPS.

Updating the renormalized operators

As previously discussed, we need different types of manipulations for updating the renormalized operators. First, all renormalized operators have to be updated as in Eq. (2.42); this process scales as $\mathcal{O}(L^2D^3)$. On the other hand, we need to take into account new contributions in the complimentary operators as was exemplified in Eq. (2.43). For this step, the most intensive part occurs when updating or constructing complementary operators of two operators. As this type of complementary operators are only needed for the longest side of the MPS chain we can choose to construct them from scratch as in Eq. (2.41) when reaching the middle of the MPS chain. On the other hand, we can also start by constructing them from the beginning and update them while progressing. Both methods scale as $\mathcal{O}(L^4D^2)$ per sweep.

Updating the renormalized operators scales in total as $\mathcal{O}(L^3D^3 + L^4D^2)$ per sweep.

⁵In practice I often find it easier to think about renormalized operators as partially traced MPOs.

The two site optimization

During the optimization of a two-site (or one-site) tensor we solve Eq. (2.31) with an iterative eigensolver like, for example, the Davidson method [157]. For this, the action of the effective Hamiltonian has to be calculated iteratively. Thanks to the complementary renormalized operators, the action of the effective Hamiltonian upon the two-site tensor scales as $\mathcal{O}(L^2 D^3)$; the most intensive part is the recombination of the renormalized operators of two second-quantized operators with their respective complementary renormalized operators. Finally, when the optimized tensor is obtained it still has to be decomposed into two separate one-site tensors through the singular value decomposition. This procedure has a complexity of $\mathcal{O}(D^3)$.

In total, the two-site optimization and the subsequent decomposition has a scaling of $\mathcal{O}(L^3 D^3)$ per sweep.

Memory and disk usage

Of course, the DMRG algorithm has to keep track of all the different optimized tensors. Besides this, we need to store the renormalized operators. As we want to recycle previously calculated renormalized operators during the sweeps, one set of these operators has to be kept in memory per virtual bond. The needed memory for this is thus $\mathcal{O}(L^3 D^2)$. Not all of these renormalized operators have to be kept in RAM however. A lot of these sets can be written to disk as the calculated renormalized operators at each bond are not used until the sweep passes by said bond.

2.4. Other tensor networks

The success of the MPS for simulating one-dimensional local gapped Hamiltonians is due to its layout as a tensor network; its linear form allows a very efficient encoding for entangled states respecting the one-dimensional area law. However, one can still attempt to tackle two-dimensional systems with the linear MPS. In fact, the MPS has been extensively used on systems such as the square Heisenberg and the square Hubbard model [158–161]. To study these, the MPS can be mapped to the square lattice in a snake-like fashion. This is represented in Fig. 2.2. Comparing the geometry of the MPS and the square lattice, it is clear that the entanglement between neighboring spins in vertical direction can be adequately captured while the MPS will have much more trouble with horizontal neighbors. Since these neighboring spins are artificially distanced in the MPS, their entanglement will be much more costly to take into account. The MPS is not a natural wave function for the study of such two-dimensional systems; the right correlations can only be captured by using a sufficiently large bond dimension and by choosing the square lattice

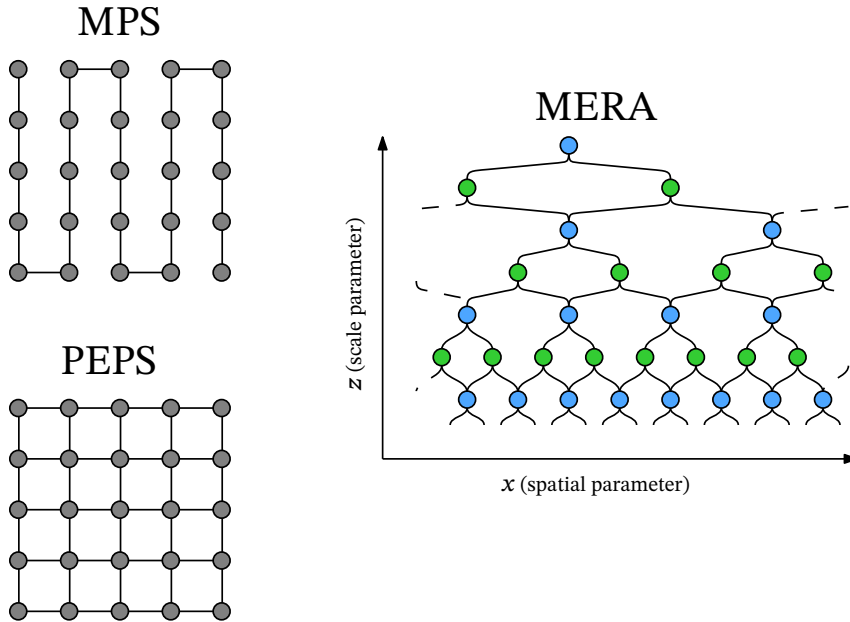


Figure 2.2: (left) Example of an MPS and a PEPS mapped to a square grid. Only the virtual bonds of the network are represented; each tensor has an additional physical bond which is not represented as to not clutter the graphic. (right) Example of a MERA. The isometries are colored blue while the disentanglers are colored green.

on a narrow strip or a cylinder.

To solve this problem the projected entangled pair states (PEPS) [162] was developed as an extension of the MPS in higher dimensions. A graphical depiction of PEPS is given in Fig. 2.2. As the PEPS mimics the geometry of the square lattice, it is much more suited to capture the correlations of that system. The better connectivity of the PEPS tensor network does come at a cost though. As the PEPS tensors in Fig. 2.2 have up to four virtual bonds, much lower bond dimensions have to be used to keep the optimization feasible. The PEPS ansatz also has loops in its network, making it harder to define a canonical form; however, it is still possible [163–165]. Furthermore, some needed tensor contractions for optimization are not exactly feasible anymore and need to be approximated.

For d -dimensional critical systems, the tensor network does not need to mimic the physical geometry of the lattice; it should mimic the so-called $(d+1)$ -dimensional *holographic* geometry where the scale parameter z is added as extra dimension. These types of tensor networks are called the multi-scale entanglement renormalization ansatz (MERA) [166, 167]. The $(1+1)$ -dimensional

case is represented in Fig. 2.2.

Other tensor networks states than the MPS have also been used in quantum chemical studies. Notable examples are the tree tensor network states (TTNS) [168–170], complete-graph tensor network states (CGTNS) [171] and self-adaptive tensor network states (SATNS) [172]. Although it is debatable if the SATNS and CGTNS should be considered tensor networks; they can be written as tensor networks, but their approximation stems from imposed correlations of the tensor elements instead of a truncated bond dimension. Although its one-dimensional nature is far from ideal save for linear molecules, the MPS is still the preferred tensor network for chemical systems. DMRG is characterized by a high efficiency and stability of the algorithm and a relative ease of implementing SU(2)-symmetry. This allows us to largely make up for the suboptimal entanglement representation by increasing the used virtual bond dimension.

It is also worth pointing out continuous versions were formulated for both the MPS and the MERA and are used for quantum fields. They are, quite fittingly, called the continuous MPS (cMPS) [173, 174] and the continuous MERA (cMERA) [175].

2.4.1. Tree Tensor Networks

The tree tensor network state is a natural extension of the MPS ansatz. While the MPS wave function can be depicted as a linear chain of tensors, the TTNS ansatz allows branching of the network. It is the most general tensor network state without any loops. By using this ansatz, a better representation of the entanglement topology of the system is expected as compared to the MPS, since component tensors can have an arbitrary order. An example of the TTNS is shown in Fig. 2.3a.

An advantage of the TTNS is that at a finite bond dimension it is able to capture algebraically decaying correlation functions. This in contrast to DMRG which is only able to represent exponentially decaying correlations [168, 170, 176]. Let us clarify this a bit; imagine we start from one central tensor and we radially expand the TTNS with a fixed coordination number z (i.e. the maximum number of virtual bonds of a tensor in the TTNS). The number of sites L in function of the number of layers Y is

$$L = 1 + z \sum_{k=1}^Y (z-1)^{k-1}, \quad (2.44)$$

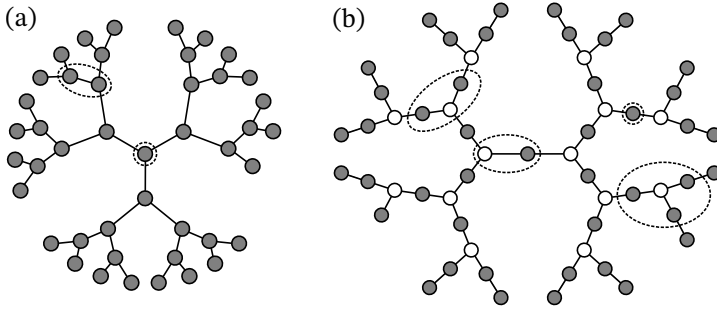


Figure 2.3: (a) A general TTNS with maximum $z = 3$ and 44 orbitals. The tensor optimized during one or two-site optimization is shown by dashed lines, and has maximally 3 or 4 virtual bond indices, respectively. (b) An example of a T3NS with 44 orbitals. The tensor optimized during one, two-, three-, or four-site optimization is shown by the dashed contours, and has maximally 3 virtual bond indices for all cases. Filled circles represent *physical tensors* and have thus an extra physical index that is not drawn here for simplicity. Empty circles represent *branching tensors*.

or

$$L = \frac{z(z-1)^Y - 2}{z-2} \quad z \geq 3 \quad (\text{TTNS}) \quad (2.45)$$

$$L = 1 + 2Y. \quad z = 2 \quad (\text{MPS}) \quad (2.46)$$

The maximal distance between two sites is given by $2Y$. From Eq. (2.45) and Eq. (2.46) follows a logarithmic scaling of the maximal distance with system size L for trees and a linear one for MPSs. Correlation functions in TTNSs with finite bond dimension decay exponentially in function of maximal distance. Hence, in function of system size an algebraic decay is obtained for $z \geq 3$ in contrast to the exponential decay for the MPS ($z = 2$) [176–178].

In quantum chemistry, the tree tensor network ansatz was first developed for the study of dynamics. Independently discovered from general tensor network methods, it is known as the multilayer multiconfiguration time-dependent Hartree theory (ML-MCTDH) [179–181]. For the study of electronic structure theory, the QC-TTNS was first used by Murg *et al.* [168, 170] for trees with arbitrary coordination number. The complexity of the algorithm as a function of the virtual dimension D is given by $\mathcal{O}(D^{x+1})$, where x is given by the coordination number of the tensor optimized at each stage. Due to this scaling, Murg *et al.* restricted themselves to a maximum coordination number of 3 in the network and to a one-site optimization scheme; this results in an optimization of $\mathcal{O}(D^4)$. Using a two-site optimization in the same tree with coordination number of 3 would involve the optimization of two-site tensors with 4 virtual bonds (as seen in Fig. 2.3a); such optimization results in an expensive $\mathcal{O}(D^5)$.

In DMRG, the usage of a two-site optimization scheme has proved to be advantageous. The two-site scheme is less prone to be stuck in local minima and an automatic redistribution of the virtual dimensions over different symmetry sectors is easily obtained through SVD [182]. In TTNS it would be opportune to also use two-site optimization. In contrast to DMRG though, two-site optimization for an arbitrary TTNS is accompanied with a heavier polynomial cost than one-site optimization as previously stated. In the work of Nakatani *et al.* [169] this problem is circumvented by introducing *half-renormalization*. In the half-renormalization step the TTNS is exactly mapped to an MPS. In this MPS, the iterative optimization step is executed at a DMRG-like cost. The mapping of the TTNS to an MPS is still expensive, though.

In the next section, we will propose the three-legged tree tensor network state (T3NS) ansatz which has considerable advantages compared to a general TTNS. I will show that it enables two-site optimization with only a prefactorial penalty in the polynomial scaling and without the need of the half-renormalization scheme.

The TTNS is not unique to quantum chemistry. Just as for the MPS, it was first studied in condensed matter physics [177, 183]. Even in machine learning the TTNS enjoys some interest [184].

2.5. The T3NS ansatz

Just as in previous works on QC-TTNS [168–170], we restrict ourselves to a maximum coordination number $z = 3$, to keep calculations feasible. A second restriction we impose, is that only tensors with $z \leq 2$ have physical indices. We call this type of tensors *physical tensors*. Tensors with $z = 3$ are called *branching tensors* and have exclusively virtual bond indices. We do not allow branching tensors to be adjacent to each other. An example of a T3NS is given in Fig. 2.3b.

The proposed ansatz enables us to go beyond one-site optimization and use two-site, three-site or even four-site optimization with the same polynomial scaling (see Fig. 2.3b). Another substantial advantage of T3NS is that every tensor has at most three different indices (one physical and two virtual for a physical tensor and three virtual for a branching tensor). This will be rather useful for the implementation of a $SU(2)$ -adapted version, which is the main subject of Chapter 3.

There are some similarities between the T3NS and the half-renormalization scheme of Nakatani and Chan [169]. The mapping of the TTNS to the MPS used for half-renormalization corresponds with a branching tensor in our T3NS formulation. In the half-renormalization scheme the branching tensor is *never directly* optimized; consequently, more sweeps are needed for conver-

gence. A one-site optimization using half-renormalization is equivalent with a one-site optimization for T3NS where the branching tensors are skipped during optimization. For a two-site optimization using half-renormalization the equivalence is less apparent; swapping of tensors in the network during the T3NS optimization would be needed for a sound comparison. Although rather trivial, it is also worth pointing out that the MPS ansatzes are a subset of the T3NS; the MPS corresponds with a T3NS devoid of branching tensors.

Our T3NS implementation is able to do two-site, three-site and four-site optimization, but in this dissertation only two-site optimization is used except if mentioned otherwise. Three- and four-site calculations have been executed but the small increase in energy accuracy did not outweigh the extra computational time needed (larger prefactor). However, the ability to do three- and four-site optimization can be useful for orbital optimizations.

2.5.1. The T3NS algorithm

Optimization for the T3NS wave function is quite equivalent with DMRG as discussed in Section 2.2.3. In a sweeping motion, one tensor or a few neighboring tensors are picked and optimized. During the sweep, the environment of the picked tensors can be chosen trivial by choosing the appropriate gauge; the absence of loops in the T3NS allows us to define a canonical form completely equivalent as the one discussed in Section 2.2.1 for the MPS.

While a sweep through the linear MPS did not leave much room for imagination, this is now different. Many ways of stepping through the network can be chosen. In this dissertation we will always use a sweep resembling a depth-first search in a tree graph; it is probably the most obvious sweep and treats all branches and tensors on equal footing. One could also opt to skip complete branches during a sweep when these are deemed converged or weakly correlated with the rest of the system. However, we did not use any sweeping schemes like this.

Just as QC-DMRG can use an MPO formulation of the quantum chemical Hamiltonian, one could opt for a tensor network operator (TNO) formulation and represent said Hamiltonian by a tree tensor network. Just as for the MPO, a first condition for an efficient TNO is a scaling of $\mathcal{O}(L^2)$ for the virtual bond dimension. Again utmost care in handling the sparsity of the TNOs is needed to obtain similar costs as with renormalized operators; failing to do so is even more catastrophic for QC-TTNS than for QC-DMRG. The methods proposed in refs. [133, 134] are not readily translatable to QC-TTNS or do not produce the same scaling as with renormalized operators. For this reason we opted for renormalized operators, just as in previous work on QC-TTNS [168–170]. As the usage of renormalized operators for the T3NS only differs minimally

with the DMRG case, we refer the reader to Section 2.3.1 for a more thorough discussion.

2

2.5.2. Resource costs of the algorithm

Let us now compare both the CPU time and the memory usage for the QC-T3NS algorithm with the QC-DMRG algorithm. The mentioned scaling in this sections is valid up to the four-site optimization.

The most time consuming part of the algorithm is the iterative execution of the action of the effective Hamiltonian on the combined tensor. Due to the usage of complementary renormalized operators, the effective Hamiltonian is constructed out of $\mathcal{O}(L^2)$ different terms for both DMRG and T3NS. However, the cost of constructing each term scales as $\mathcal{O}(D^4)$ for T3NS instead of $\mathcal{O}(D^3)$ for DMRG.

The other leading term for the computational complexity is due to updating the renormalized operators. The most intensive type of update for the renormalized operators is one unique to the T3NS algorithm and occurs when two sets of renormalized operators have to be recombined in a new set by using a branching tensor. The most intensive type of recombination occurs when a single operator in both sets has to be updated to a complimentary double operator, e.g.

$$\begin{aligned}
 & \text{Diagram of two operators } c_{i\uparrow}^\dagger \text{ and } c_{j\uparrow}^\dagger \text{ connected by a branching tensor} \\
 &= \boxed{c_{i\uparrow}^\dagger c_{j\uparrow}^\dagger} \rightarrow \sum_{ij} [V_{ijcd} - V_{ijdc}] \boxed{c_{i\uparrow}^\dagger c_{j\uparrow}^\dagger} \rightarrow \boxed{P_{d\uparrow ct}}. \quad (2.47)
 \end{aligned}$$

First, the single operators are combined to a double operator with the aid of the branching tensor; an operation with complexity $\mathcal{O}(L^2 D^4)$. In the second stage, the newly formed double operators are summed together with their potential terms into the different double complementary operators; for the worst case scenario this scales as $\mathcal{O}(L^4 D^2)$. As there are $\mathcal{O}(L)$ occasions per sweep for this, a conservative total scaling of $\mathcal{O}(L^5 D^2 + L^3 D^4)$ could be assumed; it is the scaling reported in Ref. [104]. However, only the most central branching tensors will result in the worst case scaling of $\mathcal{O}(L^4 D^2)$; an exponentially small portion of the total amount of branching tensors. A more careful analysis results in an upper bound of $\mathcal{O}(L^4 D^2 + L^3 D^4)$ for the updating of the renormalized operators. This was corrected in the appendix of Ref. [105]. A summary of the scaling of both QC-DMRG and QC-T3NS is given in Table 2.1.

At a fixed system size k and bond dimension D , the speed of a sweep is still dependent of the particular shape of the tree. While one can only make one

	QC-DMRG	QC-T3NS
CPU time:	$\mathcal{O}(L^4 D^2 + \underline{L^3 D^3})$	$\mathcal{O}(L^4 D^2 + \underline{L^3 D^4})$
Memory:	$\mathcal{O}(L^3 D^2)$	$\mathcal{O}(L^3 D^2 + LD^3)$

Table 2.1: Resource requirements of DMRG and T3NS with renormalized operators. The underlined terms correspond with the action of the effective Hamiltonian; it is the iterative portion of the sparse eigensolver.

shape of MPS-chain for a fixed system size, this is not true for trees. We will now briefly touch upon this.

2.6. Shape of the tree, choice and ordering of the orbitals

The different shapes for the tree at fixed L introduce additional freedom that is not present in DMRG. It is clear that the particular shape will influence the speed and the accuracy of the calculations as it determines the correlations captured by the network. The orbital ordering in the network is also of importance for the accuracy of the calculations. This freedom is also present in DMRG and multiple methods for ordering the orbitals exist (e.g. through use of the mutual information [170] or the exchange integral [169] have been studied). Similar methods can be used to optimize the shape of the tree. Finally, the orbital choice and orbital optimization is also of importance for TTNS and DMRG calculations. Quite some research has been done for this in DMRG [149, 185–189]. In TTNS, orbital optimization by canonical transformations has been studied and used [168].

Optimization of the shape of the network and the orbital order will be the primary subject of Chapter 5.

2.7. Fermionic Networks

A quantum chemical calculation involves fermions. This introduces extra complexity in the algorithm through the sign change of the wave function when interchanging two fermions. Historically, the most common way to tackle fermionic systems for the MPS is by mapping the fermionic creation and annihilation operators onto spin operators, known as the Jordan-Wigner transformation [190]. However, we opted for the fermionic network formalism in this work as developed by Bultinck *et al.* [191]. It provides a clear and simple way for the introduction of the fermionic signs without the need of too much bookkeeping. In this formalism, an example of an arbitrary fermionic tensor

is given by

$$\mathbf{A} = \sum_{\alpha\beta\gamma\delta\dots} A_{\alpha\beta\gamma\delta\dots} |\alpha\rangle|\beta\rangle\langle\gamma|\langle\delta|\dots \quad (2.48)$$

This is equivalent with the definition for bosonic tensors, but where $|\alpha\rangle, |\beta\rangle, \langle\gamma|, \langle\delta|, \dots$ are instead elements of the so-called super vector space V . The fermionic signs are introduced by the following canonical isomorphism

$$\begin{aligned} \mathcal{F} : V \otimes_{\mathfrak{g}} W &\rightarrow W \otimes_{\mathfrak{g}} V \\ |i\rangle \otimes_{\mathfrak{g}} |j\rangle &\rightarrow (-1)^{|i||j|} |j\rangle \otimes_{\mathfrak{g}} |i\rangle, \end{aligned} \quad (2.49)$$

where V and W are super vector spaces and $\otimes_{\mathfrak{g}}$ denotes the graded tensor product. $|i\rangle$ and $|j\rangle$ represent homogeneous basis states. A homogeneous state is characterized by a definite parity of the state, i.e. $|i|, |j| \in \{0, 1\}$. For the contraction of fermionic tensors, a second mapping is introduced:

$$\mathcal{C} : V^* \otimes_{\mathfrak{g}} V \rightarrow \mathbb{C} : \langle\psi| \otimes_{\mathfrak{g}} |\phi\rangle \rightarrow \langle\psi|\phi\rangle, \quad (2.50)$$

where V and V^* are the super vector space and its dual space, respectively.

When contracting two fermionic tensors, the states of the tensors should first be ordered appropriately through successive usage of Eq. (2.49) before using Eq. (2.50). For further details we refer the reader to Ref. [191].

2.8. Numerical results

At this point, we have all the tools for a working T3NS implementation. we will proceed by an initial comparison of the T3NS ansatz with the MPS. The results were previously published in Ref. [104] and were generated with a first test-implementation of the T3NS ansatz exploiting only the particle conservation of the chemical Hamiltonian. This is not the same implementation as the one available on github [156]; the presented timings are not representative for the latter implementation. However, general conclusions remain valid.

Having obtained a solid grasp on the workings of both the MPS and the T3NS ansatz, we will now compare them in a quantum chemical setting. We compare for several systems the energy errors and CPU times in function of the bond dimension. We study LiF and N_2 at their equilibrium bond length ($r = 3.05$ a.u. and $r = 2.118$ a.u., respectively). For LiF we also perform calculations at longer bond lengths. LiF and N_2 are two systems that don't particularly call for a tree-shaped topology representation. However, as we will show, a similar accuracy is already obtained with the T3NS at considerable lower bond dimension as compared with DMRG for both systems. This fact

gives us hope that the more complex entanglement topology will prove even more its merits in larger molecules, since the orbitals can be easier arranged in groups of highly entangled orbitals [192].

LiF and N_2 are studied for different bond dimensions by T3NS with U(1)-symmetry and DMRG with U(1)-symmetry. LiF is also studied with DMRG with SU(2)-symmetry. For both T3NS-U(1) and DMRG-U(1) we use our own implementation of the T3NS ansatz. For DMRG-SU(2) we use the ChemPS2 software program developed by S. Wouters [92, 144, 176, 193]. Both systems are popular benchmarks for methods and their ability to take strong electron correlations into account, and both systems have been studied in previous papers about QC-TTNS [169, 170].

To test the T3NS ansatz in larger systems, we perform calculations on the bis(μ -oxo) and the $\mu - \eta^2 : \eta^2$ peroxy $[Cu_2O_2]^{2-}$ isomers. We compare the energy gaps between the two isomers obtained by T3NS with previously published values [142, 149–152, 194–197]. We also compare the obtained complete active space (CAS) ground-state energies with other calculations executed in the same active space; as to keep comparisons fair.

For now, we will group orbitals belonging to the same spatial irrep as much as possible and connect the irreps in the center of the tree. As a rudimentary rule, one could assume that orbitals belonging to the same irrep are strongly correlated as only these are connected to each other through the kinetic terms of the chemical Hamiltonian. Within one irrep, the orbitals are ordered such that the orbitals closest to the Fermi level (for LiF and N_2) or the orbitals with highest single-orbital entropy (for $[Cu_2O_2]^{2-}$) are closest to the center of the tree. After experimenting with a few different orderings of the orbitals, this proved to be the most successful one. Another degree of freedom is the choice of the orbitals. In this paper, we only use Hartree-Fock orbitals. No orbital optimization is done at this point. We refer the interested reader to the supplementary material of Ref. [104] for the used T3NS geometries and orbital orderings used in these calculations.

2.8.1. Lithium fluoride

The first system studied with the new T3NS ansatz is LiF. We perform calculations at equilibrium bond length ($r = 3.05$ a.u.), at $r = 12$ a.u. where an avoided crossing occurs, and at large bond length ($r = 13.7$ a.u.). Calculations are performed in a CAS of size (6e, 25o). The atomic orbital basis from Bauschlicher and Langhoff [198] and a frozen core of 1σ , 2σ and 3σ orbitals was used. The same basis set and active space is used in Ref. [170]. Ground state energies were calculated by using T3NS-U(1), DMRG-U(1) and DMRG-SU(2) with varying bond dimensions. In the case of DMRG-SU(2), the quoted

bond dimension is the reduced one, where the additional SU(2)-symmetry is taken into account [176]. FCI energies were easily recovered through the T3NS. Accuracies in the order of $10^{-8} E_h$ were obtained for all bond lengths at $D = 400$. The accuracy of DMRG and T3NS in relation to the bond dimension is given in Fig. 2.4 for LiF in equilibrium. As expected, a lower bond dimension is needed for T3NS for a similar accuracy as in DMRG. In Fig. 2.5, the accuracy of DMRG and T3NS is given for different bond lengths at a low bond dimension ($D = 100$).

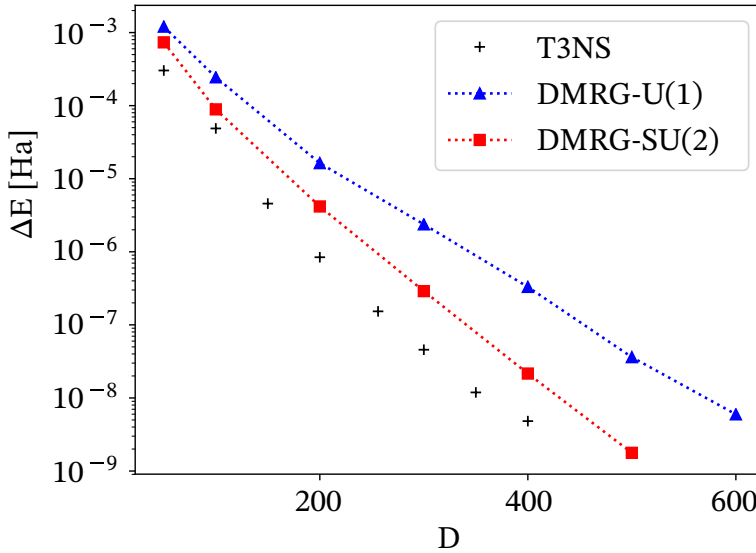


Figure 2.4: Energy difference of DMRG and T3NS calculations with respect to the FCI energy for LiF at equilibrium bond length $r = 3.05$ a.u. FCI energies are obtained from Ref. [170]. The calculations are done at different (*reduced*) bond dimensions.

	CPU time last sweep	total CPU time
T3NS[100]	96 sec	428 sec
T3NS[400]	1000 sec	2240 sec
DMRG[100]	48 sec	600 sec
DMRG[600]	640 sec	1884 sec

Table 2.2: Some timings for T3NS and DMRG calculations of LiF at equilibrium bond length. Used bond dimensions are given in square brackets. Both T3NS and DMRG are executed with our own implementation to keep comparison fair.

Lastly, some wall times for the T3NS and DMRG calculations are given in Table 2.2 for $r = 3.05$. At $D = 100$, a sweep is twice as slow in T3NS as

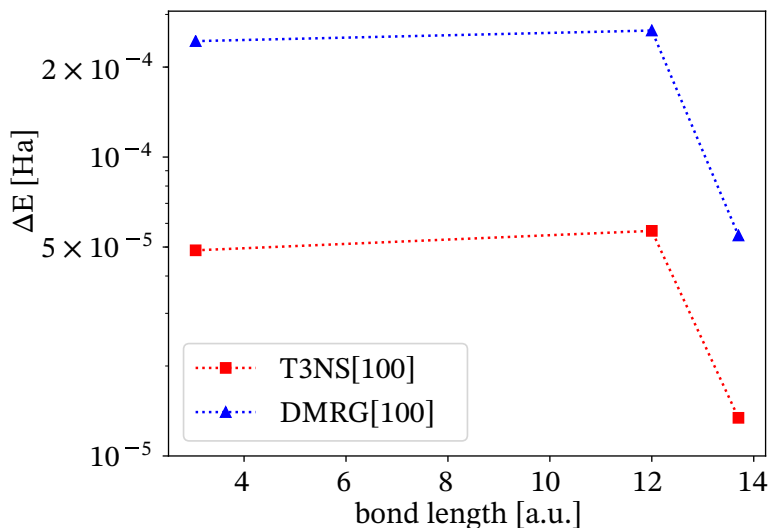


Figure 2.5: Energy difference of DMRG-U(1) and T3NS-U(1) calculations with respect to the FCI energy for LiF at bond length $r = 3.05, 12$ and 13.7 a.u. FCI energies are obtained from Ref. [170]. The calculations are done at $D = 100$ for both DMRG and T3NS.

in DMRG, as can be expected. However, less sweeps are needed until convergence which ultimately results in a faster calculation with higher accuracy. The need for fewer sweeps in T3NS is something we noticed quite consistently. For DMRG at $D = 600$ and T3NS at $D = 400$ both accuracy and total wall time are comparable. We would like to note that these remarks on timing are by no means conclusive since the speed and accuracy of both T3NS and DMRG are heavily dependent on orbital ordering and initial guess. In these calculations, a random initial guess and a rather intuitive orbital ordering was used. These remarks are merely to illustrate the competitiveness of our T3NS ansatz with DMRG.

2.8.2. The nitrogen dimer

The second benchmark system for our T3NS ansatz is the nitrogen dimer at equilibrium (bond length: 2.118 a.u.). This is a popular molecule for benchmarking methods in their ability to describe strong electron correlation accurately. Because of this it has also been discussed by Nakatani *et al.* in their TTNS paper [169]. They studied the nitrogen dimer in a cc-pVDZ basis set in a frozen core active space (10e, 26o), keeping the 1s electrons of nitrogen fixed. DMRG [199] and FCI [200] calculations have also been previously executed for this active space. In this paper we execute all-electron calculations (14e, 28) for the nitrogen dimer in a cc-pVDZ basis set and compare them with the

most accurate results obtained in Ref. [199] through coupled cluster on the SDTQPH level.

Several calculations have been executed at different bond dimensions for T3NS and DMRG with $U(1)\times U(1)$ -symmetry. The obtained energy differences with respect to CCSDTQPH [199] are given in Fig. 2.6 for bond dimensions up to 1000 for DMRG and up to 700 for T3NS. Comparable energies are obtained for T3NS at half the bond dimension needed for DMRG. This is consistent with the conclusion from the frozen core TTNS calculations in Ref. [169].

	CPU time last sweep	total CPU time
T3NS[100]	560 sec	1440 sec
T3NS[300]	4 h	17 h
T3NS[500]	16 h	96 h
T3NS[700]	66 h	237 h
DMRG[100]	160 sec	3800 sec
DMRG[500]	2050 sec	9 h
DMRG[1000]	2 h	27 h

Table 2.3: Some CPU times for T3NS and DMRG calculations of N_2 at equilibrium bond length. Both T3NS and DMRG are executed with our own implementation to keep comparison fair.

CPU times are given in Table 2.3 for T3NS- $U(1)$ and DMRG- $U(1)$ calculations. Similar conclusions can be made in comparison with LiF. For $D = 100$, T3NS-sweeps take longer than DMRG-sweeps, but the number of sweeps needed for convergence from a random initial guess is considerably lower. This ultimately results in a lower wall time. At $D = 1000$ for DMRG and $D = 500$ for T3NS, obtained accuracies are comparable. Wall times for T3NS are higher though than for DMRG, but still in the same order.

2.8.3. The bisoxo and peroxo isomer of $[Cu_2O_2]^{2-}$

As a last benchmark system, we study the bisoxo(μ -oxo) and peroxo isomers of $[Cu_2O_2]^{2-}$, and in particular their energy gap. These transition metal clusters have been studied with a wide range of ab initio methods like CASSCF, CASPT2 (complete active space self consistent field theory with perturbation theory up to second order) [194] and RASPT2 (restricted active space self consistent field theory with perturbation theory up to second order) [195]. However, the small active spaces used for CASPT2 and RASPT2 showed to be insufficient. Later on, the usage of DMRG-based methods allowed to take a considerably larger active space into account, yielding improved results [142, 149–152, 196, 197].

We will use the T3NS algorithm to treat the two isomers in a (26e, 44) ac-

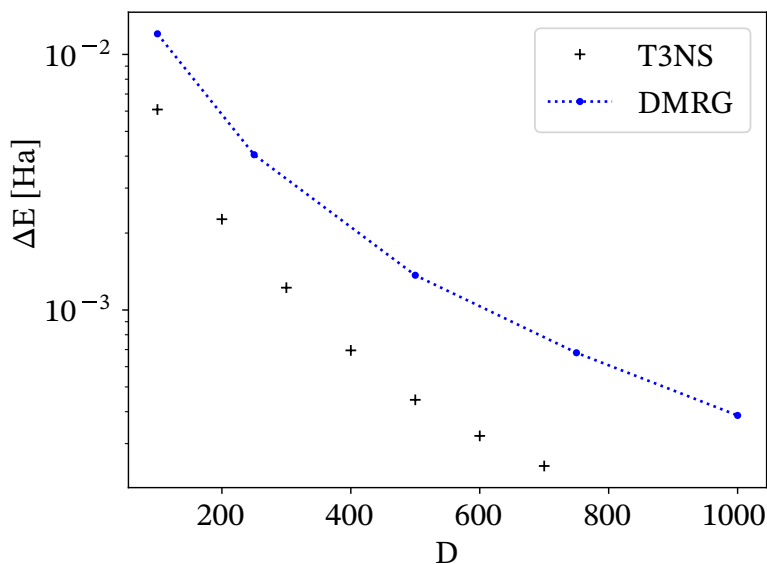


Figure 2.6: Energy difference of DMRG and T3NS calculations for the nitrogen dimer with respect to CCSDTQPH energy ($-109.282172 E_h$) [199]. The calculations are done at different bond dimensions.

tive space. The same active space and basis set is used as in Ref. [152] and [149]; the basis set is given by the Stuttgart pseudopotential and associated basis functions (ECP10MDF) [201] for Cu and the atomic natural orbital (ANO) basis set of Pierloot *et al.* for O [202]. Results are given in Table 2.4. Energies of the isomers in the used active spaces are very comparable to the ones in Ref. [149]. Furthermore, the energy gap between the two isomers are in the same region as previously executed DMRG calculations.

These results are especially promising since no advanced methods were used to augment the T3NS calculations in contrast to the previous research with DMRG for this system. At this moment our algorithm starts from a random initial guess, no effort was made in avoiding local minima and an intuitive orbital ordering was used. In contrast, previous DMRG research included the configuration interaction based dynamically extended active space (CI-DEAS) procedure [149] or perturbative noise was added to the tensors [142] to avoid local minima. More advanced orderings were also used in this previous research; orbitals were ordered by minimizing quantum entanglement using the Fiedler vector [142, 149], or a genetic algorithm [142]. Other methods used to augment the results were DMRG-SCF (self consistent field) [150], DMRG-SCF with canonical transformation theory (DMRG-SCF/CT) [150] or DMRG-CASPT2 to take dynamical correlation into account. In Ref. [149] dynamic

Ref. Method	$E_{\text{bisoxo}} [E_h]$	$E_{\text{peroxo}} [E_h]$	ΔE [kcal/mol]
[194] CASSCF(16,14)			0.2
[194] CASPT2(16,14)			1.4
[195] RASPT2(24,28)			28.7
Some previously published DMRG energies			
[151] DMRG(32,62)[2400]			35.6
[150] DMRG(28,32)[2048]-SCF/CT			27.0
[197] DMRG(32,28)[4000]			21.8
[142] DMRG(24,24)[1500]-SCF*			35.1
[142] DMRG(24,24)[1500]-CASPT2*			23.2
[152] DMRG(26,44)[800]	-541.467 79	-541.497 31	18.5
[196] DMRG(26,44)[128]	-541.473 08	-541.514 70	26.1
[149] DMRG(26,44)[256/1024/10 ⁻⁵] [†]	-541.538 53	-541.581 14	26.7
T3NS calculations			
T3NS(26,44)[50]	-541.487 73	-541.569 99	51.6
T3NS(26,44)[100]	-541.523 52	-541.571 66	30.2
T3NS(26,44)[200]	-541.532 84	-541.577 17	27.8
T3NS(26,44)[300]	-541.535 56	-541.579 66	27.7
T3NS(26,44)[500]	-541.538 20	-541.580 94	26.8

*: Bond dimensions given for these calculations are *reduced* bond dimensions.

[†]: This calculation uses the DBSS method and CI-DEAS as initialization procedure. The square brackets state that a minimum of $D = 256$ is used at every bond, the CI-DEAS procedure starts with $D = 1024$ and a maximum discarded weight of 10^{-5} is aimed for. Maximum bond dimensions around 2000 were reported for both clusters during these calculations [149].

Table 2.4: Energy gaps between the bis(μ -oxo) and $\mu - \eta^2 : \eta^2$ peroxo $[\text{Cu}_2\text{O}_2]^{2-}$ isomers from T3NS calculations of this paper and previous calculations. The energy gaps are given in kcal/mol. Ground state energies are given for the T3NS calculations and DMRG calculations from previous research using the same active space and are given in Hartree. Bond dimensions are given in square brackets.

block state selection (DBSS) was used to tune the bond dimension. Instead of a fixed bond dimension, a maximum discarded weight is used. In this way, the bond dimension at every bond is tailored to stay below this maximum discarded weight. DBSS is easily implementable once two-site optimization is used, like in our T3NS algorithm. Since we noted that the discarded weight was dominant in very few bonds while it was orders lower in other bonds, we think that DBSS can also yield a substantial improvement in the T3NS algorithm.

To check if we got stuck in local minima, the ground state wave function of the bisoxo isomer obtained through T3NS[500] was compressed to a lower bond dimension. The compressed wave function was then reoptimized at this lower bond dimension and we found a ground state energy of -541.50527 , -541.52387 and -541.53327 Hartree for $D = 50, 100$ and 200 , respectively. Comparison with the results obtained through random initialization in Table 2.4 makes the problem of local minima quite clear and shows us that preventing local minima can improve our results significantly, especially at low bond dimension.

2.9. A short recap

As the three-legged tree tensor network state (T3NS) is a direct extension of the density matrix renormalization group (DMRG), we first thoroughly explained this algorithm for the calculation of ground state wave functions in quantum chemistry; thereafter, the T3NS algorithm can be introduced with only minor changes and extensions.

Although the T3NS is computationally more expensive than DMRG, our hope is that the increased flexibility of the tensor network allows a more accurate entanglement representation at a lower number of renormalized states. As a first cautious assessment, some numerical results are given in which DMRG and T3NS are compared for a selection of chemical systems; they indicate that T3NS is at least competitive with DMRG for the given systems and does indeed give improved results at lower bond dimensions.

We started the chapter by a brief overview of the origin of DMRG and tensor networks in its whole. One can only guess if P.A.M. Dirac would still maintain his view when seeing the plethora of developments sprung from renormalization. At least, we do hope the reader does not share his opinion after reading this chapter.



3

Symmetries

*...out of love of symmetry,
just as people put two vases above a fireplace.*

Guy de Maupassant

In the previous chapter, the ansatz for the T3NS was formulated. Although the exponential wall is circumvented by using insights into entanglement, we can still do better by looking at the symmetries in the problem at hand.

Chemical systems have several symmetries. Symmetries due to particle conservation and abelian point group symmetries are easily exploitable; others, such as non-abelian point group symmetries and the total electron spin conservation, need more involved techniques. By using the Wigner-Eckart theorem we will discuss how SU(2)-symmetries can be exploited. Although briefly explained in the previous section, we will also go deeper into tackling fermionic systems.

Parts of this chapter have been published in Journal of Chemical Theory and Computation **15**, 2996-3007 (2019) [105].

Symmetries are a keystone for efficient quantum chemical calculations as they naturally restrict the search space of the algorithm to the relevant corner of the Hilbert space. Already the earliest QC-DMRG formulations leveraged the particle number and spin projection conservation of the chemical Hamiltonian to its advantage [123, 203, 204]. It did not take long before any real abelian point group symmetries of the molecule could also be exploited [185, 199]. Although of limited use for general molecules, they can provide substantial speedups when relevant.

While the abelian symmetries of the chemical Hamiltonian were quite straightforward to tackle for DMRG, it was not the case for non-abelian symmetries. The most important non-abelian symmetry for quantum chemistry is given by the total spin conservation of the electrons which is an SU(2)-symmetry. Sierra and Nishino first included SU(2)-symmetries in DMRG through the interaction round a face Hamiltonian [205]. Later on, McCulloch and Gulácsi introduced another simpler technique [206–208]. Translated to the MPS, their work assumed that each three-legged tensor in the MPS is given by a irreducible tensor operator; using the Wigner-Eckart theorem [209], these tensors can be decomposed in a reduced part and the relevant Clebsch Gordan coefficients. Being the bread-and-butter of tensor networks with non-abelian symmetries, let us formally define the Wigner-Eckart theorem for SU(2):

Theorem 3.1 (Wigner-Eckart) *When expressed in an eigenbasis of the angular momentum each tensor operator $T_M^{(J)}$ which transforms similarly as a $|JM\rangle$ angular momentum eigenstate under rotations is called an irreducible tensor. Such tensors can be decomposed as follows*

$$\langle j' m' | T_M^{(J)} | j m \rangle = \langle j' || \mathbf{T}^{(J)} || j \rangle C_{m M m'}^{j j' J} \quad (3.1)$$

where $\mathbf{T}^{(J)}$ is the reduced tensor independent of the angular momentum projections and $C_{m M m'}^{j j' J}$ are the Clebsch-Gordan coefficients.

Usage of the Wigner-Eckart theorem was soon common practice for SU(2)-adapted DMRG, either in condensed matter or quantum chemistry [148, 176, 210–213]. Also for other tensor networks methods the Wigner-Eckart theorem proved to be useful although it can be more involved [214–216]. The technique introduced by McCulloch and Gulácsi is easily extensible to other ‘multiplicity-free’ non-abelian symmetries, such as demonstrated by Sharma for the $D_{\infty h}$ point group symmetry of the C_2 dimer [217].

Weichselbaum showed that non-multiplicity-free non-abelian symmetries such as SU(3) can also be exploited by tensor networks [218]. A symmetry is called non-multiplicity-free if either the same m_z label occurs multiple times

within one irrep or when the product of two irreps results into multiple occurrences of the same irrep for the direct sum decomposition. Luckily enough, the symmetries that are commonly exploited in quantum chemistry are multiplicity-free. As such, we will only focus on these types of symmetries.

3.1. A symmetry adapted three-legged tensor

The usage of symmetries in the T3NS is facilitated by restricting to a maximum of three legs for every tensor in the wave function ansatz. It allows for a straightforward application of the Wigner-Eckart theorem for the different tensors. It also makes the treatment of both abelian and non-abelian global symmetries very analogous to the MPS, where this is already well understood [148, 176, 208, 210, 213, 219, 220].

If we want to leverage the Wigner-Eckart theorem to our benefit, we need to first express the different three-legged tensors in a basis of angular momentum eigenstates. When writing the tensors in such basis we can preserve the symmetries of the system at each tensor explicitly. Even more, the tensors are now given by irreducible tensor operators; we can then split each tensor into a so-called reduced tensor and a Clebsch-Gordan coefficient.

Let us first investigate how to write the tensors in a basis of angular momentum eigenstates.

3.1.1. Labeling the basis states

In the T3NS the index of a virtual (or physical) leg specifies the different virtual (or physical) basis states traveling through this leg. When leveraging the Wigner-Eckart theorem, these basis states of the tensors need to transform as rows of the irreducible representations of the symmetry group [176, 208, 210, 219, 220]. Each basis state $|\alpha\rangle$ (or index of a leg) can thus be labeled with the irrep and the row of the irrep according to which it transforms, i.e. $|\alpha\rangle = |\text{Irreps and rows}, \alpha'\rangle$.

The label α' is needed to discern the different basis states belonging to the same list of irreps and rows of irreps. For a physical bond in non-relativistic quantum chemistry for example we need the local basis states at every orbital. We get

$$|-\rangle = |\pi = 0, N_{\uparrow} = 0, N_{\downarrow} = 0\rangle \quad (3.2)$$

$$|\uparrow\rangle = |\pi = 1, N_{\uparrow} = 1, N_{\downarrow} = 0\rangle \quad (3.3)$$

$$|\downarrow\rangle = |\pi = 1, N_{\uparrow} = 0, N_{\downarrow} = 1\rangle \quad (3.4)$$

$$|\uparrow\downarrow\rangle = |\pi = 0, N_{\uparrow} = 1, N_{\downarrow} = 1\rangle \quad (3.5)$$

for the parity \mathbb{Z}_2 (π), and two $U(1)$ -symmetries for both the spin up and down (N_\uparrow, N_\downarrow), or

$$|-\rangle = |\pi = 0, I = I_0, N = 0, j = 0, m = 0\rangle \quad (3.6)$$

$$|\uparrow\rangle = |\pi = 1, I = I_k, N = 1, j = 1/2, m = 1/2\rangle \quad (3.7)$$

$$|\downarrow\rangle = |\pi = 1, I = I_k, N = 1, j = 1/2, m = -1/2\rangle \quad (3.8)$$

$$|\uparrow\downarrow\rangle = |\pi = 0, I = I_0, N = 2, j = 0, m = 0\rangle \quad (3.9)$$

for \mathbb{Z}_2 (π), $U(1)$ (N), $SU(2)$ (j, m) and the real abelian point group symmetries (I). The labels $\pi, N_\uparrow, N_\downarrow, N, j$ and I represent irreps of the different symmetries, while m labels the row of irrep j . I_k is the point group irrep of the orbital. A double occupation results in the trivial point group irrep I_0 since $I_k \otimes I_k = I_0$ for real abelian point group symmetries. In this example, $SU(2)$ is the only symmetry that needs an extra label for the row since the other symmetries have one-dimensional representations.

We omitted the label α' , since the irreps and rows already uniquely label the local physical basis states. However for the labeling of the virtual basis states the label α' is still needed.

For calculations with fermions we utilize fermionic tensor networks. For every performed permutation or contraction the fermionic signs are calculated by looking at the parities of the different basis states [104, 191]. One could note that labeling the parity π is redundant since it is already fixed by the total number of particles in the state. However, we chose to keep explicitly track of the parity of the states. In this way, we can separate the fermionic sign handling completely from the particle numbers in the basis states. This allows a more modular implementation of the different symmetries and the fermionic signs.

3.1.2. Reduced tensors

Using the labeling discussed in the previous section, we can write every three-legged tensor in the wave function ansatz as shown in Eq. (3.11). Due to the Wigner-Eckart theorem, the irreducible tensor A can be rewritten as a reduced tensor \tilde{A}^r multiplied with the Clebsch-Gordan coefficients of the different symmetries. This is shown in Eq. (3.12). These coefficients are Kronecker deltas for the real abelian symmetries and are the well-known Clebsch-Gordan coefficients for the recoupling of spins for $SU(2)$.

We can relate the Clebsch-Gordan coefficients with the Wigner $3j$ -symbols by using their relation:

$$\langle j_1 m_1 j_2 m_2 | j_3 m_3 \rangle = \sqrt{2j_3 + 1} (-1)^{j_1 - j_2 + m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}. \quad (3.10)$$

$$\mathbf{A} = \sum_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} A_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} |\mathbf{x}_1 \alpha_1\rangle |\mathbf{x}_2 \alpha_2\rangle \langle \mathbf{x}_3 \alpha_3| \quad (3.11)$$

$$= \sum_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} \delta_{I_1 \otimes I_2, I_3} \delta_{\pi_1 + \pi_2 \bmod 2, \pi_3} \delta_{N_1 + N_2, N_3} \langle j_1 m_1 j_2 m_2 | j_3 m_3 \rangle \tilde{A}_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3}^r |\mathbf{x}_1 \alpha_1\rangle |\mathbf{x}_2 \alpha_2\rangle \langle \mathbf{x}_3 \alpha_3| \quad (3.12)$$

$$= \sum_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} \delta_I \delta_\pi \delta_N \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} (-1)^{j_3 - m_3} A_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3}^r |\mathbf{x}_1 \alpha_1\rangle |\mathbf{x}_2 \alpha_2\rangle \langle \mathbf{x}_3 \alpha_3| \quad (3.13)$$

$$= \sum_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} A_{\mathbf{x}_1; \mathbf{x}_2; \mathbf{x}_3}^s A_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3}^r |\mathbf{x}_1 \alpha_1\rangle |\mathbf{x}_2 \alpha_2\rangle \langle \mathbf{x}_3 \alpha_3| \quad (3.14)$$

$$= \sum_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} \underbrace{\begin{array}{c} \downarrow \mathbf{x}_2 \alpha_2 \\ \textcircled{A} \\ \leftarrow \mathbf{x}_1 \alpha_1 \quad \rightarrow \mathbf{x}_3 \alpha_3 \end{array}}_{\text{irreducible tensor } A} |\mathbf{x}_1 \alpha_1\rangle |\mathbf{x}_2 \alpha_2\rangle \langle \mathbf{x}_3 \alpha_3| = \sum_{\mathbf{x}_1 \alpha_1; \mathbf{x}_2 \alpha_2; \mathbf{x}_3 \alpha_3} \underbrace{\begin{array}{c} \downarrow \mathbf{x}_2 \alpha_2 \\ \textcircled{A^r} \\ \leftarrow \mathbf{x}_1 \alpha_1 \quad \rightarrow \mathbf{x}_3 \alpha_3 \end{array}}_{\text{reduced tensor } A^r} \underbrace{\begin{array}{c} \downarrow \mathbf{x}_2 \\ \textcircled{A^s} \\ \leftarrow \mathbf{x}_1 \quad \rightarrow \mathbf{x}_3 \end{array}}_{\text{symmetry tensor } A^s} |\mathbf{x}_1 \alpha_1\rangle |\mathbf{x}_2 \alpha_2\rangle \langle \mathbf{x}_3 \alpha_3| \quad (3.15)$$

where

$$\mathbf{x}_i = \pi_i I_i N_i j_i m_i$$

$$x_i = \pi_i I_i N_i j_i$$

$$\delta_I = \delta_{I_1 \otimes I_2, I_3}$$

$$\delta_N = \delta_{N_1 + N_2, N_3}$$

$$\delta_\pi = \delta_{\pi_1 + \pi_2 \bmod 2, \pi_3}$$

$$[j_i] = \sqrt{2j_i + 1}$$

$$\begin{array}{c} \downarrow \mathbf{x}_2 \\ \textcircled{A} \\ \leftarrow \mathbf{x}_1 \quad \rightarrow \mathbf{x}_3 \end{array} = \delta_{I_1 \otimes I_2, I_3} \delta_{\pi_1 + \pi_2 \bmod 2, \pi_3} \delta_{N_1 + N_2, N_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} (-1)^{j_3 - m_3}$$

In Eq. (3.13), the Clebsch-Gordan coefficients are replaced by said symbols. The $3j$ -symbols are, although rather equivalent to the Clebsch-Gordan coefficients, preferred as they are more symmetric. They are, for example, up to a phase invariant under permutations of the columns.

Several shorthand notations are introduced. \mathbf{x}_i is shorthand notation for the full labeling of irreps and rows of irreps, while x_i is the labeling of only the irreps and not the rows (m) of the basis state. Eq. (3.13) expresses the tensor \mathbf{A} in terms of a *reduced tensor* A^r independent of the labels of the rows of irreps in \mathbf{x}_i and a *symmetry tensor* A^s . This symmetry tensor consists of the remaining terms in Eq. (3.13). It contains the complete dependency of A on the rows of irreps and is completely independent of the labels α_i . Shorthand notations are also introduced for the Kronecker deltas and for the square root of the multiplicity of j_i ($[j_i] = \sqrt{2j_i + 1}$).

A phase is absorbed in A^r (hence the transition from \tilde{A}^r to A^r). The absorbed phase is real, which ensures that the elements of the reduced tensors can also be kept real for typical quantum chemical problems. Here, we also opted to absorb the $[j_i]$ term into the reduced tensor.

Graphical depiction

In Eq. (3.15) we introduce a graphical depiction of Eq. (3.13). Its introduction will facilitate further arithmetic with the symmetry adapted tensors. Let us elaborate on the graphical depiction.

On the left hand side a representation of the irreducible tensor A is given. The incoming and outgoing edges represent respectively the ket and bra states in Eq. (3.11). The circle itself represents the rank-three tensor with its different tensor elements.

On the right hand side of the equation, we introduce the $SU(2)$ -adapted graphical representation of the same tensor \mathbf{A} . The hexagon represents the reduced tensor A^r . As discussed previously, it is independent of the labeling of the rows of irreps (m). The symmetry tensor A^s is represented by a vertex with three edges. The symmetry tensor is constituted by the Wigner $3j$ -symbol, the different Kronecker deltas and an extra phase; it encodes the dependencies of the different elements in a tensor due to our imposed global symmetries.

The direction of the edges fixes the sign of m_i in the Wigner sign. An incoming edge corresponds with m_i , while an outgoing edge is $-m_i$. There is also a phase $(-1)^{j_i - m_i}$ associated with every outgoing edge of a vertex. The directed edges are also needed for a correct treatment of the fermionic signs [104, 191]; outgoing and incoming edges correspond with bra and ket states, respectively.

There is also an arrow present in the symmetry tensor arcing over the different edges. This is needed to fix the order of the bra and kets in Eq. (3.13); it

runs from the first state to the last state. Due to its fermionic nature, the tensor is not invariant under permutation of these bras and kets; the fixing of the order in our graphical depiction is needed. The arrow also represents the cyclic invariance of the $3j$ -symbol. Concerning the $SU(2)$ -symmetry, a phase has to be included when changing the direction of the arrow but not when rotating the arrow, i.e.

$$\begin{array}{c} \text{X}_2 \\ | \\ \text{x}_1 \text{---} \text{---} \text{x}_3 \end{array} = \begin{array}{c} \text{X}_2 \\ | \\ \text{x}_1 \text{---} \text{---} \text{x}_3 \end{array} = \begin{array}{c} \text{X}_2 \\ | \\ \text{x}_1 \text{---} \text{---} \text{x}_3 \end{array} = (-1)^{j_1+j_2+j_3} \begin{array}{c} \text{X}_2 \\ | \\ \text{x}_1 \text{---} \text{---} \text{x}_3 \end{array}. \quad (3.16)$$

This symmetry of the $3j$ -symbol is independent of the direction of the edges. The reader should also keep in mind that we only concentrated on the $SU(2)$ -symmetry in Eq. (3.16); the (possible) fermionic character of the tensor was disregarded. More properties of the Wigner symbols are given in Appendix B, both algebraically and graphically. We will extensively refer to properties mentioned in this appendix when needed in the following derivations.

Differences with source code

This convention differs with the one mentioned in Ref. [105] and used in the source code [156]. In these works, we opted to not absorb the $[j_i]$ terms into the reduced tensor and the phase $(-1)^{j_i-m_i}$ were not necessarily associated with the outgoing edges of vertices. Instead we explicitly kept track of these factors by introducing additional marks in the graphical representations of the tensors.

When refraining to absorb $[j_i]$, a straightforward connection is found between the reduced and irreducible tensors. The notion of a QR decomposition, singular value decomposition and orthogonal tensors are nicely equivalent for both types of tensors without the need of rescaling the tensor elements. This merit motivated us initially to keep explicitly track of the $[j_i]$ factors.

However, in the following sections we will absorb this factor and we will need to rescale the tensor elements appropriately from time to time. Working with absorbed $[j_i]$ factors is less error-prone and it is easier to formulate general rules-of-thumb for the contraction and manipulations of tensors. A simpler convention with less degrees of freedom is quite often just easier to work with. Because of this, we believe now that absorbing the factor is the better choice, certainly when a lot of different kind of contractions are needed.

Keeping explicitly track of the phase $(-1)^{j_i-m_i}$ allowed removing a few extra phases for intermediate results; we consider it now to be a quite a bit of a folly to keep track of yet another extra term for the removal of a few phases.

Bear in mind that although the idea is quite the same as in Ref. [105], the mentioned prefactors for tensor contractions are not the same; people wanting

to work directly with the source code in [156] are referred to [105] for the used conventions and the needed prefactors. People who are aiming to implement their own T3NS code are encouraged to consider the convention used in this dissertation.

Sparsity and compression of the tensors

The symmetry tensor encodes a lot of sparsity since it consists out of several Kronecker deltas and a Wigner $3j$ -symbol that respects the triangle inequality $|j_1 - j_2| \leq j_3 \leq j_1 + j_2$. It is easy to see which elements of the reduced tensor A^r do not contribute to the full tensor \mathbf{A} . Such index combinations can be set to zero without repercussions; in the next parts these tensor elements will always be assumed zero.

Next to sparsity, the symmetry tensor also compresses the data. Different indices belonging to the same multiplet in \mathbf{A} are only represented by a single index in A^r as it is independent of the rows of the irreps m . The interrelation between these different indices of the same multiplet is completely captured by the symmetry tensor instead.

3.2. Building the wave function

Only with the right tools, one can successfully build; irrespective of how fine the bricks are. Hence, let's start with clarifying tensor contractions for symmetry adapted tensors followed by proposing a canonical form. Afterwards, we can move on to actually building the T3NS wave function from the building blocks given on page 59.

3.2.1. A tensor contraction

Although tensor contractions become more efficient when exploiting the right symmetries, it is also a more elaborate endeavor than with plain tensors. Let us investigate an exemplary contraction of three tensors \mathbf{A} , \mathbf{B} and \mathbf{C} to $\mathbf{D} = \mathcal{C}(\mathbf{ABC})$. Let say that these tensors are for example given by

$$\mathbf{A} = \begin{array}{c} |x_1\alpha_1 \\ \textcircled{A} \\ x_a\alpha_a \quad x_2\alpha_2 \end{array} \quad |x_a\alpha_a\rangle|x_1\alpha_1\rangle\langle x_2\alpha_2| \quad (3.17)$$

$$= \begin{array}{c} |x_1\alpha_1 \\ \textcircled{A^r} \\ x_a\alpha_a \quad x_2\alpha_2 \end{array} \quad \begin{array}{c} \downarrow x_1 \\ \leftarrow x_a \quad \rightarrow x_2 \end{array} \quad (3.18)$$

$$\mathbf{B} = \begin{array}{c} |x_b\alpha_b \\ \textcircled{B} \\ x_1\alpha_1 \quad x_3\alpha_3 \end{array} \quad \langle x_1\alpha_1 || x_b\alpha_b \rangle \langle x_3\alpha_3 | \quad (3.19)$$

$$= \begin{array}{c} |x_b\alpha_b \\ \textcircled{B^r} \\ x_1\alpha_1 \quad x_3\alpha_3 \end{array} \quad \begin{array}{c} \downarrow x_b \\ \leftarrow x_1 \quad \rightarrow x_3 \end{array} \quad (3.20)$$

$$C = \frac{\begin{array}{c} |x_3\alpha_3 \\ \textcircled{C} \\ x_2\alpha_2 \quad x_c\alpha_c \end{array}}{|x_2\alpha_2\rangle|x_3\alpha_3\rangle\langle x_c\alpha_c|} \quad (3.21)$$

$$= \frac{\begin{array}{c} |x_3\alpha_3 \\ \textcircled{C^r} \\ x_2\alpha_2 \quad x_c\alpha_c \end{array}}{x_2 \quad x_c} \quad (3.22)$$

Note we have left out the different summations over the loose edges and the bra and ket states for each reduced-symmetry tensor decomposition as to not clutter the graphical notations. The order of bra and ket states, which is of major importance for the fermionic sign, is fixed by the arching arrow in the symmetry tensor.

For the irreducible tensors, contracting is rather straightforward; even the order of contracting the tensors can be easily optimized¹. The contraction is given by

$$\frac{\begin{array}{c} |x_b\alpha_b \\ \textcircled{D} \\ x_a\alpha_a \quad x_c\alpha_c \end{array}}{|x_a\alpha_a\rangle|x_b\alpha_b\rangle\langle x_c\alpha_c|} = \begin{array}{c} |x_b\alpha_b \\ \textcircled{B} \\ x_1\alpha_1 \quad x_3\alpha_3 \\ \textcircled{A} \quad \textcircled{C} \\ x_a\alpha_a \quad x_2\alpha_2 \quad x_c\alpha_c \end{array} \quad (3.23)$$

$\underbrace{|x_a\alpha_a\rangle|x_1\alpha_1\rangle\langle x_2\alpha_2|}_{\text{States of A}} \underbrace{\langle x_1\alpha_1||x_b\alpha_b\rangle\langle x_3\alpha_3|}_{\text{States of B}} \underbrace{|x_2\alpha_2\rangle|x_3\alpha_3\rangle\langle x_c\alpha_c|}_{\text{States of C}}$

or

$$\frac{\begin{array}{c} |x_b\alpha_b \\ \textcircled{D} \\ x_a\alpha_a \quad x_c\alpha_c \end{array}}{|x_a\alpha_a\rangle|x_b\alpha_b\rangle\langle x_c\alpha_c|} = \underbrace{(-1)^{\pi_1}}_{\text{fermionic sign}} \begin{array}{c} |x_b\alpha_b \\ \textcircled{B} \\ x_1\alpha_1 \quad x_3\alpha_3 \\ \textcircled{A} \quad \textcircled{C} \\ x_a\alpha_a \quad x_2\alpha_2 \quad x_c\alpha_c \end{array} \quad (3.24)$$

where the fermionic nature of the tensors is taken into account as explained in Section 2.7 and connected edges represent summations over said indices. It is useful to note that each tensor on its whole has a trivial parity, i.e. the total parity of all the edges of a tensor is $\pi_{\text{tot}} = 0$; by construction, each tensor transforms trivially under the parity symmetry. As a result, it does not matter in which order the states of the different tensors are concatenated in Eq. (3.23).

Instead of using the irreducible tensors directly, we can also use their symmetry-adapted decomposition into an irreducible tensor and a symmetry tensor. Now, the contraction boils down to two separate contractions; one for

¹At least, this is feasible for contractions consisting of a small set of tensors; this problem is, at its root, NP-hard.

each type of tensor, i.e.

$$\frac{x_b \alpha_b}{x_a \alpha_a} \underset{x_c \alpha_c}{D} = (-1)^{\pi_1}$$

These contractions are not independent. They are intertwined by the shared labels for the irreps (x_i). However, only the reduced tensors are dependent of α_i and only the symmetry tensors are dependent of the labels for the rows of the irreps m_i . This offers us an opportunity for simplification as the summation over m_i and α_i can be isolated to the symmetry tensors and the reduced tensors, respectively.

At this point an expression for a possible reduced tensor D^r and symmetry tensor D^s is obtained, i.e.

$$D^r = (-1)^{\pi_1} \frac{x_b \alpha_b}{x_a \alpha_a} \underset{x_c \alpha_c}{A^r} \underset{x_2 \alpha_2}{B^r} \underset{x_3 \alpha_3}{C^r}$$

$$D^s = \text{Diagram of symmetry tensor } D^s$$

Although we can calculate the reduced tensor through a straightforward contraction over the connected edges $x_1 \alpha_1$, $x_2 \alpha_2$ and $x_3 \alpha_3$, the expression for the symmetry tensor looks a bit unwieldy. Luckily enough, using the arithmetic properties of the Wigner symbols, of which the used ones are listed in Appendix B, the given symmetry tensor in Eq. (3.27) can be further simplified. As it is our first encounter which such operations we will do it quite verbosely this time:

$$D^s = \text{Diagram of symmetry tensor } D^s$$

Through usage of the reflection property:

$$= (-1)^{j_1+j_a+j_2}$$

Through usage of the 6j-swap of x_a and x_3 :

$$= \sum_{j_2'} [j_2']^2 (-1)^{2j_2+j_1+j_2'-j_3} \begin{Bmatrix} j_a & j_c & j_2' \\ j_3 & j_1 & j_2 \end{Bmatrix}$$

Through usage of the reflection and reversal property:

$$= \sum_{j_2'} [j_2']^2 (-1)^{2j_2+j_1+j_3-j_b} \begin{Bmatrix} j_a & j_c & j_2' \\ j_3 & j_1 & j_2 \end{Bmatrix}$$

Through usage of the orthogonality property:

$$= \sum_{j_2'} \delta_{x_2' x_b} (-1)^{2j_2+j_2'-j_b} \begin{Bmatrix} j_a & j_c & j_2' \\ j_3 & j_1 & j_2 \end{Bmatrix}$$

$$= (-1)^{2j_2} \begin{Bmatrix} j_a & j_c & j_b \\ j_3 & j_1 & j_2 \end{Bmatrix}$$

Through usage of the reflection property:

$$= (-1)^{2j_2+j_a+j_b+j_c} \begin{Bmatrix} j_a & j_c & j_b \\ j_3 & j_1 & j_2 \end{Bmatrix}$$

We have obtained a more compact and manageable symmetry tensor for \mathbf{D} with an additional prefactor; this prefactor can easily be absorbed in the reduced tensor. We finally obtain

$$\overline{x_a \alpha_a} \circlearrowleft \begin{matrix} | \\ x_b \alpha_b \\ | \end{matrix} \overline{x_c \alpha_c} = \overline{x_a \alpha_a} \circlearrowright \begin{matrix} | \\ x_b \alpha_b \\ | \end{matrix} \overline{x_c \alpha_c} \begin{matrix} \downarrow x_b \\ \rightarrow x_c \end{matrix} \quad (3.28)$$

with

$$\begin{aligned}
 \begin{array}{c} |x_b\alpha_b \\ \text{---} \text{D}^r \text{---} \\ x_a\alpha_a \quad x_c\alpha_c \end{array} &= \sum_{x_1\alpha_1; x_2\alpha_2; x_3\alpha_3} \begin{array}{c} |x_b\alpha_b \\ \text{---} \text{B}^r \text{---} \\ x_1\alpha_1 \quad x_3\alpha_3 \\ \text{---} \text{A}^r \text{---} \quad \text{---} \text{C}^r \text{---} \\ x_a\alpha_a \quad x_2\alpha_2 \quad x_c\alpha_c \end{array} \\
 &\times \underbrace{(-1)^{\pi_1}}_{\text{fermionic origin}} \underbrace{(-1)^{2j_2+j_a+j_b+j_c} \begin{Bmatrix} j_a & j_c & j_b \\ j_3 & j_1 & j_2 \end{Bmatrix}}_{\text{SU}(2) \text{ origin}}
 \end{aligned} \tag{3.29}$$

The summation over the contracted indices is explicitly given here to clarify that the fermionic and spin prefactors are also included in this summation. We notice that the m_i indices are left out of the final summation; they were used for the simplification of the symmetry tensor in Eq. (3.27).

Let us briefly recap how we have performed a tensor contraction for symmetry adapted tensors. First the fermionic sign can be calculated by swapping and contracting the different states up to the desired result. The order of said states is dictated by the arching arrows around the vertices in the symmetry tensors. Secondly, the reduced tensors and symmetry tensors form a separate network that can be contracted. The symmetry tensors can be simplified by contracting over the m_i indices and once the desired form is obtained, the reduced tensors can be contracted while taking into account the correct prefactors that arise from deforming the symmetry tensors.

This example does not only serves an elucidating purpose. We hope the reader also recognizes the clarity of a graphical representation opposed to the long-winded stream of indices that would have been needed in the algebraic equations, certainly since we avoided the explicit form of the symmetry tensors as in Eq. (3.13).

3.2.2. The canonical form

As discussed in Section 2.2.1.2, it is not only possible but also advantageous to define a canonical form for our wave function ansatz. In this canonical form one tensor in the network is chosen as orthogonality center; the currently optimized tensor is generally the most suitable candidate. Other tensors are orthogonal with respect to contraction over all bonds but the one leading to the orthogonality center. Calculating the overlap of the tensor network with itself now simplifies to a complete contraction of the orthogonality center as we have made the environment trivial.

In this section, a canonical form is proposed for the symmetry adapted fermionic tensors. Although it is quite straightforward to define these orthog-

onal tensors there is a small caveat concerning the fermionic character of the tensors. Due to this, we will discuss all the different cases of orthogonality and also the orthogonality center. It is also a good opportunity to further familiarize us with the graphical notations.

Orthogonal for contracting over x_1 and x_2

Let us propose a reduced tensor and a symmetry tensor for \mathbf{A} which is orthogonal to a contraction over index 1 and 2. We will also propose a reduced tensor and a symmetry tensor for the adjoint \mathbf{A}^+ , i.e. the corresponding tensor for the bra wave function.

$$\mathbf{A} = \begin{array}{c} |x_2\alpha_2 \\ \text{---} \langle A^r \text{---} \\ x_1\alpha_1 \quad x_3\alpha_3 \end{array} \begin{array}{c} \rightarrow \mathbf{x}_2 \\ \mathbf{x}_1 \quad \mathbf{x}_3 \end{array} \quad (3.30)$$

$$\mathbf{A}^+ = (-1)^{j_1+j_2-j_3} \begin{array}{c} x_1\alpha_1 \quad x_3\alpha_3 \\ \text{---} \langle \bar{A}^r \text{---} \\ |x_2\alpha_2 \end{array} \begin{array}{c} \mathbf{x}_1 \quad \mathbf{x}_3 \\ \rightarrow \mathbf{x}_2 \end{array} \quad (3.31)$$

Let us determine the conditions for orthogonality where \bar{A}^r is the complex conjugate of A^r :

$$\mathcal{C}(\mathbf{A}\mathbf{A}^+) = (-1)^{j_1+j_2-j_3} \begin{array}{c} \langle \bar{A}^r \text{---} x_3'\alpha_3' \\ |x_1\alpha_1 \quad x_2\alpha_2 \\ \text{---} \langle A^r \text{---} \\ x_3\alpha_3 \end{array} \begin{array}{c} \mathbf{x}_1 \quad \mathbf{x}_3' \\ \mathbf{x}_2 \quad \mathbf{x}_3 \end{array} \quad (3.32)$$

$$= [j_3]^{-2} \delta_{x_3, x_3'} \begin{array}{c} \langle \bar{A}^r \text{---} x_3'\alpha_3' \\ |x_1\alpha_1 \quad x_2\alpha_2 \\ \text{---} \langle A^r \text{---} \\ x_3\alpha_3 \end{array} \begin{array}{c} \mathbf{x}_3 \end{array} \quad (3.33)$$

It is clear that the reduced tensors need to fulfill

$$\begin{array}{c} \langle \bar{A}^r \text{---} x_3'\alpha_3' \\ |x_1\alpha_1 \quad x_2\alpha_2 \\ \text{---} \langle A^r \text{---} \\ x_3\alpha_3 \end{array} = \delta_{\alpha_3, \alpha_3'} [j_3]^2 \begin{array}{c} \mathbf{x}_3 \end{array} \quad (3.34)$$

Pay close attention to the labels of the loose edges in the left hand side of the equation. The label for the irreps x_3 are the same but the labels for the different states α_3 and α_3' are different. This adjusted orthogonality condition for the reduced tensors is, together with the orthogonality property of the symmetry tensors, enough to impose orthogonality of \mathbf{A} ; indeed, we obtain

$$\mathcal{C}(\mathbf{A}\mathbf{A}^+) = \begin{array}{c} \langle \bar{A}^r \text{---} \mathbf{x}_3 \\ \mathbf{x}_3 \end{array} = \sum_{x_3\alpha_3} |\mathbf{x}_3\alpha_3\rangle \langle \mathbf{x}_3\alpha_3|. \quad (3.35)$$

When comparing Eq. (3.40) with Eq. (3.31) we notice the introduction of an extra phase $(-1)^{\pi_2}$ for the former adjoint; this is needed for imposing orthogonality due to the fermionic nature of the tensors. The adjoint tensor \mathbf{A}^+ is now equal to the hermitian conjugate of \mathbf{A} up to the phase $(-1)^{\pi_2}$; this does not pose a problem as long the contraction over all adjoint tensors of the network results in the expected bra wave function; I refer the reader to Section 3.2.3 for the proof. By first contracting over the indexes m_2 and m_3 , and using the orthogonality property of the symmetry tensors and the orthogonality of the reduced tensors in Eq. (3.41) we obtain

$$\mathcal{C}(\mathbf{A}\mathbf{A}^+) = (-1)^{\pi_2}(-1)^{j_1+j_2-j_3} \begin{array}{c} x_1'\alpha_1' \\ \overline{A^r} \\ x_2\alpha_2 \\ x_3\alpha_3 \\ A^r \\ x_1\alpha_1 \end{array} \begin{array}{c} \overline{x_1'} \\ x_2 \\ x_3 \end{array} \quad (3.42)$$

$$= \begin{array}{c} |x_1\alpha_1\rangle |x_2\alpha_2\rangle |x_3\alpha_3\rangle |x_3\alpha_3\rangle |x_2\alpha_2\rangle |x_1\alpha_1\rangle \\ \overbrace{\quad}^{x_1\alpha_1} \quad \overbrace{\quad}^{x_1\alpha_1} \quad |x_1\alpha_1\rangle \langle x_1\alpha_1| \end{array} \quad (3.43)$$

$$= \sum_{x_1\alpha_1} |x_1\alpha_1\rangle \langle x_1\alpha_1|. \quad (3.44)$$

We have written out the states explicitly as to show why the factor $(-1)^{\pi_2}$ is needed.

Orthogonal for contracting over x_1 and x_3

Let us also propose a form for the reduced and symmetry tensor when the \mathbf{A} is orthogonal with respect to contractions over index 1 and 3. Just as for the previous cases, the reduced tensor is again suitably orthogonalized on its own. The reduced tensor of the adjoint is again given by the hermitian conjugate with an extra phase due to its fermionic character. The proposed forms are

$$\mathbf{A} = \begin{array}{c} x_2\alpha_2 \\ A^r \\ x_1\alpha_1 \quad x_3\alpha_3 \end{array} \begin{array}{c} \overline{x_1} \\ \overline{x_3} \end{array} \quad (3.45)$$

$$\mathbf{A}^+ = (-1)^{\pi_1}(-1)^{j_1+j_2-j_3} \begin{array}{c} x_1\alpha_1 \\ \overline{A^r} \\ x_3\alpha_3 \\ x_2\alpha_2 \end{array} \begin{array}{c} \overline{x_1} \\ \overline{x_3} \end{array} \quad (3.46)$$

By again using the orthogonality property of the symmetry tensors, we obtain the desired result for the given propositions, i.e.

$$\mathcal{C}(\mathbf{A}\mathbf{A}^+) = \sum_{x_2\alpha_2} |x_2\alpha_2\rangle \langle x_2\alpha_2|. \quad (3.47)$$

The orthogonality center

At this point, we only have to define the final building block for the canonical form, the orthogonality center. This special tensor is not in an orthogonal form and dictates how the other tensors in the network will be orthogonalized; every tensor is orthogonalized with respect to a full contraction except over the edge leading to the orthogonality center in the network. Again, we propose a reduced tensor and symmetry tensor for the orthogonality center and its adjoint:

$$\mathbf{A} = \begin{array}{c} |x_2\alpha_2 \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ x_1\alpha_1 \text{---} x_3\alpha_3 \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ |x_2\alpha_2 \end{array} \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{array} \quad (3.48)$$

$$\mathbf{A}^+ = (-1)^{\pi_3} (-1)^{j_1 + j_2 - j_3} \begin{array}{c} |x_2\alpha_2 \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ x_1\alpha_1 \text{---} x_3\alpha_3 \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ |x_2\alpha_2 \end{array} \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{array} \quad (3.49)$$

An extra phase has to be introduced once again due to the fermionic nature of the tensors. Evaluating the norm of this tensor network now boils down to a complete contraction of the orthogonality center with its adjoint; the environment is trivial through orthogonalization. We obtain

$$\mathcal{C}(\mathbf{A}\mathbf{A}^+) = (-1)^{\pi_3} (-1)^{j_1 + j_2 - j_3} \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{array} \quad (3.50)$$

$$\begin{aligned} & |\mathbf{x}_3\alpha_3\rangle \langle \mathbf{x}_2\alpha_2| \langle \mathbf{x}_1\alpha_1| |\mathbf{x}_1\alpha_1\rangle |\mathbf{x}_2\alpha_2\rangle \langle \mathbf{x}_3\alpha_3| \\ &= \sum_{x_1\alpha_1; x_2\alpha_2; x_3\alpha_3} |A_{x_1\alpha_1; x_2\alpha_2; x_3\alpha_3}^r|^2 = \|\vec{A}\|^2 = \langle \Psi | \Psi \rangle, \end{aligned} \quad (3.51)$$

where \vec{A} are the elements of the reduced tensor flattened into a vector. Bear again in mind that reduced tensor elements that do not contribute to the full tensor are assumed to be zero. Such tensor elements are elements of which the indices x_1 , x_2 and x_3 do not respect the triangle inequality of the SU(2)-symmetry or the addition rules for the abelian symmetries.

3.2.3. The complete ansatz

To construct the wave function, the tensors in the network should be contracted. Let us give an example where a T3NS is formed by contracting four branching tensors (white) and nine physical ones (gray); the network of irreducible tensors decomposes into an equivalent network of reduced tensors and

symmetry tensors, i.e.

$$(3.52)$$

We can bring the given tensor network in its canonical form by orthogonalizing the reduced tensors through QR decompositions² towards the dedicated orthogonality center.

For clarity, all the physical bonds are labeled with \mathbf{x}_p in the symmetry tensor network. Looking closer at the network, we note that the physical tensors at the border of the network have one dangling uncontracted virtual bond; this is needed since we impose that all physical tensors have two virtual bonds; it allows us to work with symmetry-invariant tensors. Most of the extra edges merely introduce the trivial vacuum state $|\mathbf{x}_{\text{vac}}\rangle$; for one edge it corresponds with an outgoing state $|\mathbf{x}_{\text{last}}\rangle$. Changing the allowed quantum numbers for this state, lets us manage the quantum numbers of the targeted ground state. This is equivalent with the singlet-embedding strategy introduced by Sharma and Chan for spin-adapted DMRG [148]. Normally, this last edge will have a *reduced* bond dimension of one; this does not have to be the case, for example when targeting multiple states at once or a state that is not an eigenstate of the used symmetries, see Chapter 4.

Completely contracting this wave function results in an approximation of the needed exact wave function, i.e.

$$\mathcal{C}(\text{T3NS}) = |-\rangle|-\rangle \dots |-\rangle \sum_{ijkl\dots} C_{ijkl\dots} |ijkl\dots\rangle \langle \mathbf{x}_{\text{last}}| \quad (3.53)$$

$$= \underbrace{|-\rangle|-\rangle \dots |-\rangle}_{\text{vacuum states}} |\Psi\rangle \langle \mathbf{x}_{\text{last}}|, \quad (3.54)$$

where $C_{ijkl\dots}$ is the approximate FCI tensor; it is obtained by completely contracting the reduced and symmetry tensors. As a next step, the bra wave function is built from the adjoint tensors as defined in Section 3.2.2. Now, when taking the extra fermionic phases into account for the defined adjoints, we ob-

²See Section C.1 for SU(2)-adapted QR decomposition.

tain

$$\mathcal{E}(\text{T3NS}^+) = |\mathbf{x}_{\text{last}}\rangle \sum_{ijkl\dots} (-1)^{\pi_{\text{last}}} C_{ijkl\dots}^* \langle ijkl\dots | \langle - | \langle - | \dots \langle - | \quad (3.55)$$

$$= |\mathbf{x}_{\text{last}}\rangle \langle \Psi | \langle - | \langle - | \dots \langle - | (-1)^{\pi_{\text{last}}} \quad (3.56)$$

where π_{last} is the parity of the complete wave function $|\Psi\rangle$ and $|\mathbf{x}_{\text{last}}\rangle$. This results into

$$\mathcal{E}(\text{T3NS} \cdot \text{T3NS}^+) = \langle \Psi | \Psi \rangle. \quad (3.57)$$

When contracting the complete tensor network and when contracting its adjoint, we get the ket and the bra of the wave function with their CI coefficients complex conjugates, as should be. The contraction of the adjoint gives rise to an extra phase due to the fermionic nature. Without it, the complete contraction of the tensor network and its adjoint would result in the square of the norm up to a phase.

The power of using symmetry adapted tensors is that the original network of irreducible tensors factorizes into two networks with the same shape. One network consists of reduced tensors; it covers the complete dependency of the labels α_i and is also dependent of the labeling of irreps. The other network is built from symmetry tensors; the dependency of the wave function on the rows of irreps (m) is completely captured by this network and it is independent of α_i .

3.3. Building the chemical Hamiltonian

As a next step, let us discuss how to write the Hamiltonian of our system into a suitable form. The non-relativistic quantum chemical Hamiltonian is given by

$$H = E_0 + \sum_{ij} T_{ij} \sum_{\sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + \frac{1}{2} \sum_{ijkl} V_{ijkl} \sum_{\sigma\tau} c_{i\sigma}^{\dagger} c_{j\tau}^{\dagger} c_{l\tau} c_{k\sigma}, \quad (3.58)$$

where i, j, k and l are indices for the different spatial orbitals and σ, τ represent the spin degree of freedom (\uparrow or \downarrow). For the four-point interactions this separates into the following cases:

$$i = j, k = l : V_{iikk} c_{i\uparrow}^{\dagger} c_{i\downarrow}^{\dagger} c_{k\downarrow} c_{k\uparrow} \quad (3.59)$$

$$i = j, k < l : V_{iikl} (c_{i\uparrow}^{\dagger} c_{i\downarrow}^{\dagger} c_{l\downarrow} c_{k\uparrow} + c_{i\uparrow}^{\dagger} c_{i\downarrow}^{\dagger} c_{k\downarrow} c_{l\uparrow}), \quad (3.60)$$

$$i < j, k = l : V_{ijkk} (c_{i\uparrow}^{\dagger} c_{j\downarrow}^{\dagger} c_{k\downarrow} c_{k\uparrow} + c_{j\uparrow}^{\dagger} c_{i\downarrow}^{\dagger} c_{k\downarrow} c_{k\uparrow}), \quad (3.61)$$

$$i < j, k < l : V_{ijkl} (c_{i\uparrow}^{\dagger} c_{j\uparrow}^{\dagger} c_{l\uparrow} c_{k\uparrow} + c_{i\uparrow}^{\dagger} c_{j\downarrow}^{\dagger} c_{l\downarrow} c_{k\uparrow} + c_{i\downarrow}^{\dagger} c_{j\uparrow}^{\dagger} c_{l\uparrow} c_{k\downarrow} + c_{i\downarrow}^{\dagger} c_{j\downarrow}^{\dagger} c_{l\downarrow} c_{k\downarrow}) \\ + (k \leftrightarrow l), \quad (3.62)$$

where we only used $V_{ijkl} = V_{jilk}$, which is due to the electrons being identical particles. The complete Hamiltonian can be constructed by summing these terms for all possible (i, j, k, l) -combinations.

The creation and annihilation operators $c_{k\sigma}^\dagger$ and $c_{k\sigma}$ do not yet transform according to the rows of the SU(2)-irreps. As is well known, an additional phase has to be introduced [176]. One possible transformation is given by

$$\tilde{c}_{k\sigma}^\dagger = c_{k\sigma}^\dagger \tag{3.63}$$

$$\tilde{c}_{k\sigma} = (-1)^{\frac{1}{2} + \sigma} c_{k-\sigma}. \tag{3.64}$$

In this way, we can again split off the Clebsch-Gordan coefficients into a symmetry tensor. As an example, this results for an annihilation operator on orbital k in

$$\left\{ \begin{array}{l} \sigma = -1/2 : c_{k\uparrow} \\ \sigma = 1/2 : c_{k\downarrow} \end{array} \right\} = \begin{array}{c} x_{k'} \\ \text{---} \tilde{c}_k \text{---} \\ x_k \end{array} \begin{array}{c} \mathbf{x}_{k'} \\ \downarrow \\ \mathbf{x}_k \end{array} \rightarrow \mathbf{x}_{kk'}. \tag{3.65}$$

Here, \mathbf{x}_k and $\mathbf{x}_{k'}$ are the local physical states for orbital k . The third edge $\mathbf{x}_{kk'} = (\pi = 1, I = I_k, n = -1, j = 1/2, j_z = \sigma)$ serves the purpose of correctly coupling different operators; its usage will be clarified later on. The bonds do not need an extra α label; the local physical states are already uniquely labeled by \mathbf{x}_k . The creation operators and other one-orbital operators have completely equivalent decompositions into a reduced tensor and a symmetry tensor. The elements of the reduced tensors can be easily calculated by hand.

Having obtained a symmetry adapted description of the creation and annihilation operators, let us try to write the terms in Eq. (3.62) in an SU(2)-adapted way, i.e.

$$\tag{3.66}$$

where

$$\begin{aligned} \mathbf{x}_{ii'} &= (1, I_i, 1, 1/2, \sigma_i) & \mathbf{x}_{jj'} &= (1, I_j, 1, 1/2, \sigma_j) \\ \mathbf{x}_{kk'} &= (1, I_k, -1, 1/2, \sigma_k) & \mathbf{x}_{ll'} &= (1, I_l, -1, 1/2, \sigma_l) \\ \mathbf{x}_{ij} &= (0, I_i \otimes I_j, 2, J, -M) & \mathbf{x}_{kl} &= (0, I_k \otimes I_l, -2, J, M). \end{aligned}$$

Here we have first fused the creation and annihilation operators separately together; thereafter, the two groups are fused to the trivial irrep as the term in Eq. (3.62) transforms trivially under said symmetries.

Fusing the creation operators and the annihilation operators results into

$$(c_i^\dagger c_j^\dagger)_{JM} = \begin{array}{c} \begin{array}{c} x_{j'} \\ \tilde{c}_j^\dagger \\ x_j \end{array} \begin{array}{c} x_{jj'} \\ \tilde{c}_i^\dagger \\ x_{ii'} \\ x_i \end{array} \begin{array}{c} [J] \\ x_{ij} \end{array} \begin{array}{c} \mathbf{x}_{j'} \\ \mathbf{x}_j \\ \mathbf{x}_{jj'} \\ \mathbf{x}_{ii'} \\ \mathbf{x}_i \end{array} \begin{array}{c} \mathbf{x}_{ij} \end{array} \end{array} \quad (3.67)$$

$$(c_k c_l)_{JM} = \begin{array}{c} \begin{array}{c} x_k \\ \tilde{c}_k \\ x_{kk'} \\ x_{kl} \\ [J] \\ x_{ll'} \\ x_l \\ x_{l'} \end{array} \begin{array}{c} x_{kk'} \\ x_{kl} \\ x_{ll'} \\ x_l \end{array} \begin{array}{c} \mathbf{x}_k \\ \mathbf{x}_{kk'} \\ \mathbf{x}_{kl} \\ \mathbf{x}_{ll'} \\ \mathbf{x}_l \end{array} \begin{array}{c} \mathbf{x}_k \\ \mathbf{x}_{kk'} \\ \mathbf{x}_{ll'} \\ \mathbf{x}_l \end{array} \end{array} \quad (3.68)$$

where J can be 0 or 1; two fermionic spin-half operators can fuse together to a singlet or a triplet. One can work out that these fused operators are given by

$$(c_i^\dagger c_j^\dagger)_{0M} = \left\{ \frac{1}{\sqrt{2}} (c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger - c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger) \right\} \quad (3.69)$$

$$(c_i^\dagger c_j^\dagger)_{1M} = \left\{ -c_{i\uparrow}^\dagger c_{j\uparrow}^\dagger, -\frac{1}{\sqrt{2}} (c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger + c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger), -c_{i\downarrow}^\dagger c_{j\downarrow}^\dagger \right\} \quad (3.70)$$

$$(c_k c_l)_{0M} = \left\{ -\frac{1}{\sqrt{2}} (c_{k\downarrow} c_{l\uparrow} - c_{k\uparrow} c_{l\downarrow}) \right\} \quad (3.71)$$

$$(c_k c_l)_{1M} = \left\{ -c_{k\downarrow} c_{l\downarrow}, \frac{1}{\sqrt{2}} (c_{k\downarrow} c_{l\uparrow} + c_{k\uparrow} c_{l\downarrow}), -c_{k\uparrow} c_{l\uparrow} \right\}, \quad (3.72)$$

We now fuse these creation and annihilation operators as depicted in Eq. (3.66); the eventually goal is to obtain Eq. (3.62) and determining the value of the reduced tensor F_J . Fusing the singlet operators results into

$$-\frac{1}{2} (c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger c_{k\downarrow} c_{l\uparrow} + c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger c_{l\downarrow} c_{k\uparrow} + c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger c_{l\uparrow} c_{k\downarrow} + c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger c_{k\uparrow} c_{l\downarrow}), \quad (3.73)$$

while for the triplet operators we obtain

$$\begin{aligned}
 & -\frac{1}{\sqrt{3}}c_{i\uparrow}^\dagger c_{j\uparrow}^\dagger c_{l\uparrow} c_{k\uparrow} - \frac{1}{\sqrt{3}}c_{i\downarrow}^\dagger c_{j\downarrow}^\dagger c_{l\downarrow} c_{k\downarrow} \\
 & + \frac{1}{2\sqrt{3}}(c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger c_{k\downarrow} c_{l\uparrow} - c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger c_{l\downarrow} c_{k\uparrow} - c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger c_{l\uparrow} c_{k\downarrow} + c_{i\downarrow}^\dagger c_{j\uparrow}^\dagger c_{k\uparrow} c_{l\downarrow}).
 \end{aligned} \tag{3.74}$$

By choosing

$$F_J = -[J](V_{ijkl} + (-1)^J V_{ijlk}) \tag{3.75}$$

the contraction results in Eq. (3.62).

While we have obtained a SU(2)-invariant formulation of the Hamiltonian, this is not yet sufficient for evaluating the energy efficiently; the underlying tensor network structure of the Hamiltonian term in Eq. (3.66) should also reflect the structure of the T3NS wave function. Obtaining an arbitrary T3NS structure out of this is possible through manipulation of the network. First, the order of the operators (e.g. switching c_i^\dagger and c_k) can be changed by using the reflection and 9j-swap property of the Wigner symbols and taking an appropriate fermionic sign into account. For example, we can swap \tilde{c}_j^\dagger and \tilde{c}_k as follows:

(3.76)

Note that the structure is still the same but we have swapped the order of k and j . For this to be the same value as Eq. (3.66) we have to change F_J to \tilde{F}_J . Let us calculate this new value \tilde{F}_J ; due to the 9j-swap property and the inclusion

of the fermionic phase we have

$$\begin{aligned}
 & \text{Network} = \sum_J -[\tilde{J}]^4 \begin{Bmatrix} i & j & J \\ k & l & J \\ \tilde{j} & \tilde{j} & 0 \end{Bmatrix} \\
 & \text{Network} = \sum_J -[J][\tilde{J}](-1)^{1+J+J'} \begin{Bmatrix} 1/2 & 1/2 & J \\ 1/2 & 1/2 & J \\ 1/2 & 1/2 & J' \end{Bmatrix} F_J
 \end{aligned}
 \tag{3.77}$$

where J is the angular momentum of states \mathbf{x}_{ij} and \mathbf{x}_{kl} and \tilde{J} is the angular momentum of states \mathbf{x}_{ik} and \mathbf{x}_{jl} . With this, we can derive \tilde{F}_J :

$$[\tilde{J}]^2 \tilde{F}_J = \sum_{J=0,1} -[J]^2 [\tilde{J}]^4 \begin{Bmatrix} 1/2 & 1/2 & J \\ 1/2 & 1/2 & J \\ \tilde{j} & \tilde{j} & 0 \end{Bmatrix} F_J \tag{3.78}$$

$$\tilde{F}_J = \sum_J -[J][\tilde{J}](-1)^{1+J+J'} \begin{Bmatrix} 1/2 & 1/2 & J \\ 1/2 & 1/2 & J' \end{Bmatrix} F_J \tag{3.79}$$

This finally gives us

$$\tilde{F}_J = [J](\delta_{J,0} 2V_{ijkl} + (-1)^{\delta_{J,0}} V_{ijlk}), \tag{3.80}$$

where we have used the following property of the 9j-symbol [62]

$$\begin{Bmatrix} j_1 & j_2 & j \\ j_3 & j_4 & j \\ j' & j' & 0 \end{Bmatrix} = \frac{(-1)^{j_2+j_3+j+j'}}{[j][j']} \begin{Bmatrix} j_1 & j_2 & j \\ j_4 & j_3 & j' \end{Bmatrix}. \tag{3.81}$$

One can also insert identities³ into the network of Eq. (3.66). These transformations suffice to change the network to an arbitrary T3NS network. A

³See Section C.2

possible example of this could be

$$E_j \times \text{[Tensor Network]} \text{ , (3.82)}$$

here we have deformed the networks in Eq. (3.66) to mimic the networks of the T3NS wave function given in Eq. (3.52). The blue tensors are the needed operators working in on both the bra and ket T3NS. The tensors denoted with $\mathbb{1}$ are identities. The exact value of their reduced tensors depend on the coupling-order; for more info see Section C.2.

The takeaway is that both the wave function ansatz and the terms of the Hamiltonian can be represented by the same tensor network shape and both are factorized into a reduced tensor network and a symmetry tensor network.

3.4. Optimizing the wave function

In Section 3.2.2 we obtained the tools to build an $SU(2)$ -symmetric wave function in canonical form. In Section 3.3 we learned how the chemical Hamiltonian can be written in a $SU(2)$ -symmetric form and can be transformed to mimic the T3NS geometry; this allows us to efficiently evaluate $\langle \Psi | \hat{H} | \Psi \rangle$. As discussed in Section 2.5.1, the optimization of a T3NS occurs in a similar way as for DMRG, i.e. we sweep through the network and optimize only a few tensors at a time. During this local optimization of the network, the effect of the Hamiltonian on the other tensors (the environment) can be efficiently captured by renormalized operators [123, 151, 176, 204].

In the following sections, the optimization of the $SU(2)$ -adapted T3NS will be discussed by considering a T3NS as given in Fig. 3.1. Both wave function and Hamiltonian have been molded into a $SU(2)$ -friendly form and have the same network geometry (as discussed in previous sections). As to not overwhelm the reader with too much diagrams, we will only highlight a few of the possible needed contractions in the main body of this dissertation. The execution of the effective Hamiltonian on a branching and neighboring physical

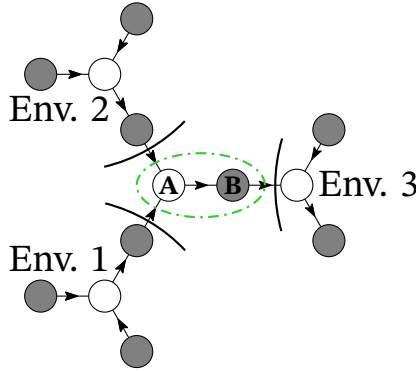


Figure 3.1: Graphical depiction of a tensor network with nine physical tensors (gray) and three branching tensors (white). An exemplary current state for the optimization is indicated. The tensors **A** and **B** are optimized; the three relevant environments for these tensors are also indicated. We also indicate the ‘flow’ imposed by the symmetry tensors. The physical edges of the physical tensors are left out.

site (a *T3NS-like* optimization) and appending a site-operator to a preexisting renormalized operator are shown in the parts to come. For some other types of updates of the renormalized operators and for the execution of the effective Hamiltonian on two physical tensors (a *DMRG-like* optimization), we redirect the reader to Section C.3 in Appendix C; although we imagine these extra diagrams will only give extra insight to people aiming for their own implementation.

Without further ado, let us discuss the two-site optimization of the T3NS shown in Fig. 3.1.

3.4.1. The two-site tensor

When performing the given two-site optimization, a good initial guess for the optimal two-site tensor is given by the contraction of the two current one-site tensors; i.e. in Fig. 3.1, the branching tensor **A** and the physical tensor **B** should be contracted. We work with a suitably canonicalized tensor network; either **A** or **B** is the orthogonality center and the other is appropriately orthogonal. Contracting the two tensors results in

$$\mathbf{M} = \mathcal{C}(\mathbf{AB}) = \begin{array}{c} x_2\alpha_2 \\ \diagdown \\ \text{A}^r \text{---} \text{B}^r \\ \diagup \\ x_1\alpha_1 \end{array} \begin{array}{c} x_p \\ | \\ x_{12}\alpha_{12} \\ | \\ x_3\alpha_3 \end{array} \begin{array}{c} x_2 \\ \diagdown \\ x_{12} \\ \diagup \\ x_1 \end{array} \begin{array}{c} x_p \\ \diagdown \\ x_{12} \\ \diagup \\ x_3 \end{array} \quad (3.83)$$

$$= \begin{array}{c} x_2\alpha_2 \\ \diagdown \\ \text{M}_{x_{12}}^r \\ \diagup \\ x_1\alpha_1 \end{array} \begin{array}{c} x_p \\ | \\ x_{12}\alpha_{12} \\ | \\ x_3\alpha_3 \end{array} \begin{array}{c} x_2 \\ \diagdown \\ x_{12} \\ \diagup \\ x_1 \end{array} \begin{array}{c} x_p \\ \diagdown \\ x_{12} \\ \diagup \\ x_3 \end{array} . \quad (3.84)$$

While we could simplify the symmetry tensor further in the example in Section 3.2.1, this is not the case in this example. We also notice that the tensor is not only dependent of the indices of the loose edges but also x_{12} . We will call x_{12} an *internal index or edge* of the tensor \mathbf{M} . This internal edge is needed to label how the two *outer* edges \mathbf{x}_1 and \mathbf{x}_2 couple; due to the non-abelian SU(2) symmetry, the coupling can have several result for a given \mathbf{x}_1 and \mathbf{x}_2 .

We have conveniently labeled the loose edges of the two-site tensor in Eq. (3.84). The local physical edge is given by \mathbf{x}_p and edges 1, 2 and 3 couple to the respective environments in Fig. 3.1. Let us further the discussion by investigating the renormalized operators for said environments.

3.4.2. The SU(2)-adapted renormalized operators

Renormalized operators are needed for efficiently treating the effect of the Hamiltonian on the environment. For the T3NS we discern two *types* of renormalized operators with each two *flavors*.

The environment of a renormalized operator can be either ‘upstream’ or ‘downstream’ to the current optimized tensors. In Fig. 3.1 for example, the renormalized operators of environment 1 and 2 are upstream while the one for environment 3 is downstream. These are the two possible flavors of the renormalized operators.

For each environment, we could contract the corresponding parts of the Hamiltonian as given in Eq. (3.82) and the bra and ket wave function as given in Eq. (3.52). In Fig. 3.1 this would mean we precontract the evaluation of $\langle \Psi | \hat{H} | \Psi \rangle$ up to the black curve for each environment. Let us call these renormalized operators *minimal*. For environment 3, we could decide to also append an operator acting upon the physical basis of tensor \mathbf{B} . Let us call this second type *extended*.

For the two types and flavors of the renormalized operators we propose a symmetry tensor structure and reduced tensor as given in Table 3.1. Just as in Section 3.4.1, we notice the need of internal edges for the extended renormalized operators; two internal edges are needed to uniquely label the couplings. Only the states labeled by $\mathbf{x}_1\alpha_1$, $\mathbf{x}_1'\alpha_1'$, $\mathbf{x}_3\alpha_3$ and $\mathbf{x}_3'\alpha_3'$ need the extra α label; these are the virtual states of the T3NS. The others are either local physical states \mathbf{x}_p and $\mathbf{x}_{p'}$ which do not need the extra label or auxiliary virtual ‘MPO’-like states \mathbf{x}_H . These govern the correct coupling of the renormalized operators to form a Hamiltonian which transforms trivially under the symmetries.

The extended renormalized operator is *not* strictly needed for the algorithm; one could refrain from appending the site-operator to the renormalized operator; instead the action of the effective Hamiltonian upon the two-site tensor \mathbf{M} could be calculated by first contracting the site operator with \mathbf{M} followed

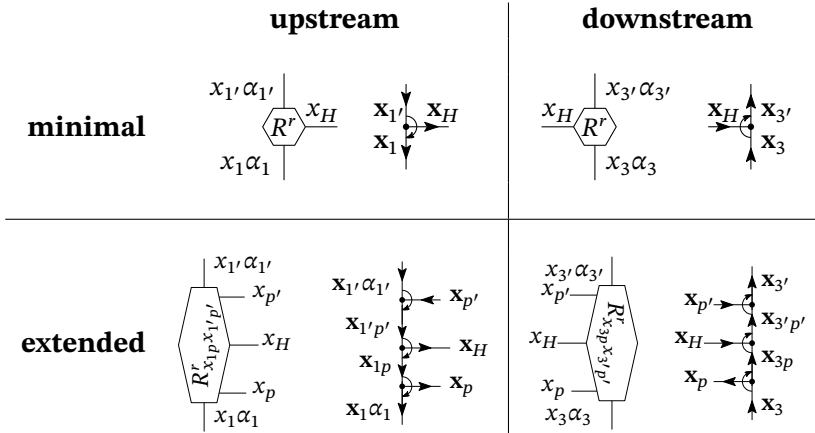


Table 3.1: Proposed forms of the symmetry tensors and reduced tensors of the different types and flavors of the renormalized operators.

by a contraction of the minimal renormalized operator. We have opted however to explicitly append the site-operators to the renormalized operators as it allows a straightforward extension of the algorithm to three- and four-site optimizations. No new deformations of the symmetry tensors have to be implemented for the action of the effective Hamiltonian. As an extra advantage, fewer high Wigner symbols are needed during the iterative action of the effective Hamiltonian; for a *DMRG-like* optimization there is even a complete absence of Wigner symbols (see Section C.3.3). In short, we move some of the deformations of the tensors that normally happen in the iterative execution of the effective Hamiltonian to a preprocessing step.

Appending a site operator to the renormalized operator

In Eq. (3.82) an example of a Hamiltonian term deformed to mimic the T3NS wave function was shown. Now, assume we already have the minimal renormalized operator \mathbf{R} for environment 3 stored in memory and an appropriate update of this renormalized operator is for example to append site-operator c_p ; this is an operator acting upon the physical basis of \mathbf{B} , as given in Eq. (3.82). After appending this site-operator, we can deform the renormalized operator to the proposed extended form in Table 3.1, i.e.

$$\begin{array}{c}
 x_{p'} \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_{pp'} \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_H \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_H \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_3 \alpha_3
 \end{array}
 \begin{array}{c}
 x_p \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_{pp'} \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_H \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_H \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_3
 \end{array}
 =
 \begin{array}{c}
 x_3', \alpha_3' \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_H \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_p \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_3 \alpha_3
 \end{array}
 \begin{array}{c}
 x_3' \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_{p'} \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_H \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_{p'} \\
 \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} \\
 x_3
 \end{array}
 \quad (3.85)$$

We can transform the symmetry tensor as follows:

$$\begin{array}{c} x_{p'} \\ \swarrow \\ x_{pp'} \\ \swarrow \\ x_H \end{array} \begin{array}{c} x_p \\ \searrow \\ x_H' \\ \searrow \\ x_3 \end{array} = \begin{array}{c} x_p \\ \swarrow \\ x_{pp'} \\ \swarrow \\ x_H \end{array} \begin{array}{c} x_{pp'} \\ \searrow \\ x_H' \\ \searrow \\ x_3 \end{array} \begin{array}{c} x_3' \\ \downarrow \end{array} \quad (3.86)$$

by using the reversal property we obtain

$$= (-1)^{2j_{H'}} \begin{array}{c} x_p \\ \swarrow \\ x_{pp'} \\ \swarrow \\ x_H \end{array} \begin{array}{c} x_{pp'} \\ \searrow \\ x_H' \\ \searrow \\ x_3 \end{array} \begin{array}{c} x_3' \\ \uparrow \end{array} \quad (3.87)$$

followed by the 9j-swap property

$$= (-1)^{2j_{H'}} [j_{1p}]^2 [j_{1'p'}]^2 \left\{ \begin{array}{ccc} j_p & j_{p'} & j_{pp'} \\ j_3 & j_{3'} & j_{H'} \\ j_{3p} & j_{3'p'} & j_H \end{array} \right\} \begin{array}{c} x_3' \\ \swarrow \\ x_3 \\ \swarrow \\ x_H \end{array} \begin{array}{c} x_p \\ \searrow \\ x_H \end{array} \quad (3.88)$$

$$= (-1)^{2j_{H'}} [j_{1p}]^2 [j_{1'p'}]^2 \left\{ \begin{array}{ccc} j_p & j_{p'} & j_{pp'} \\ j_3 & j_{3'} & j_{H'} \\ j_{3p} & j_{3'p'} & j_H \end{array} \right\} \begin{array}{c} x_p' \\ \swarrow \\ x_H \\ \swarrow \\ x_3 \end{array} \begin{array}{c} x_3' \\ \searrow \\ x_3' \\ \searrow \\ x_3 \end{array} \quad (3.89)$$

and by using reflection and reversal we finally obtain

$$= (-1)^{2j_{H'} + j_3 + j_p + j_{3'p'} + 2j_{3p'}} [j_{1p}]^2 [j_{1'p'}]^2 \left\{ \begin{array}{ccc} j_p & j_{p'} & j_{pp'} \\ j_3 & j_{3'} & j_{H'} \\ j_{3p} & j_{3'p'} & j_H \end{array} \right\} \begin{array}{c} x_p' \\ \swarrow \\ x_H \\ \swarrow \\ x_3 \end{array} \begin{array}{c} x_3' \\ \searrow \\ x_3' \\ \searrow \\ x_3 \end{array} \quad (3.90)$$

The fermionic phase is easily determined by writing out the states explicitly in the order imposed by the arching arrows around the vertices:

$$\langle x_p | \langle x_{pp'} | | x_{p'} \rangle | x_H \rangle | x_{pp'} \rangle \langle x_{H'} | | x_3 \rangle | x_{H'} \rangle \langle x_3' | = (-1)^{\pi_{pp'} + \pi_H + \pi_{p'} + \pi_{H'}} | x_3 \rangle \langle x_p | \langle x_{3p} | | x_{3p} \rangle | x_H \rangle \langle x_{3'p'} | | x_{3'p'} \rangle | x_{p'} \rangle \langle x_3' | \quad (3.91)$$

To be concise, we have dropped the labels α_i . The results for the $SU(2)$ -symmetry and the fermionic sign result together in

$$\begin{array}{c} x_{3'}\alpha_{3'} \\ x_{p'} \\ \vdots \\ x_H \\ \vdots \\ x_p \\ x_3\alpha_3 \end{array} \left[\begin{array}{c} R^r \\ x_{3p}x_{3p'} \\ x_{3p'} \end{array} \right] = \sum_{j_{pp'}, j_{H'}} (-1)^{\pi_{pp'}\pi_H + \pi_{p'}\pi_{H'}} (-1)^{2j_{H'} + j_3 + j_p + j_{3p'} + 2j_{3p'}} \quad (3.92)$$

$$[j_{1p}]^2 [j_{1p'}]^2 \left\{ \begin{array}{ccc} j_p & j_{p'} & j_{pp'} \\ j_3 & j_{3'} & j_{H'} \\ j_{3p} & j_{3p'} & j_H \end{array} \right\} \begin{array}{c} x_{p'} \begin{array}{c} \bar{c}_p \\ \vdots \\ x_p \end{array} \\ x_{pp'} \begin{array}{c} \bar{1} \\ \vdots \\ x_H \end{array} \\ x_H \begin{array}{c} \bar{R}^r \\ \vdots \\ x_3\alpha_3 \end{array} \end{array} \begin{array}{c} x_3\alpha_{3'} \\ \vdots \\ x_3\alpha_3 \end{array}$$

3.4.3. The effective Hamiltonian: the T3NS case

Let us move on to the optimization of the two-site optimization of \mathbf{M} given in Eq. (3.84) where a branching and a physical tensor were contracted. One of the original three-legged tensors was the orthogonality center of the wave function while the other one was appropriately orthogonalized; as such, the environment of the two-site tensor is trivial. The multi-site tensor has three loose virtual legs; three sets of renormalized operators are needed for the formation of the effective Hamiltonian, i.e.

$$\mathbf{H} = \sum_{\{\mathbf{RST}\}} \mathcal{C}(\mathbf{RSMT}). \quad (3.93)$$

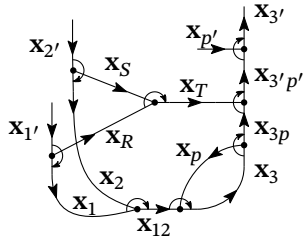
In the above equation we need to sum over all different valid combinations of renormalized operators, i.e. $\{\mathbf{RST}\}$ where \mathbf{R} and \mathbf{S} are minimal renormalized operators and \mathbf{T} is an extended renormalized operator. Let's investigate the action of one such set of renormalized operators on the multi-site tensor \mathbf{M} :

$$\mathbf{H}_{\text{part}} = \mathcal{C}(\mathbf{RSMT}) \quad (3.94)$$

$$\begin{array}{c} x_{2'}\alpha_{2'} \\ \vdots \\ x_{p'} \\ x_{3'}\alpha_{3'} \end{array} \left[\begin{array}{c} S^r \\ x_S \\ \vdots \\ x_T \\ \vdots \\ T^r \\ x_{3p}x_{3p'} \\ x_{3p'} \end{array} \right] = \begin{array}{c} x_{1'}\alpha_{1'} \\ \vdots \\ x_2\alpha_2 \\ \vdots \\ x_1\alpha_1 \\ \vdots \\ M^r_{x_{12}} \\ \vdots \\ x_3\alpha_3 \end{array} \begin{array}{c} R^r \\ \vdots \\ C \\ \vdots \\ x_p \end{array} \begin{array}{c} x_{3'}\alpha_{3'} \\ \vdots \\ x_{p'} \\ x_{3p'} \\ \vdots \\ x_{3p} \\ \vdots \\ x_3 \end{array} \quad (3.95)$$

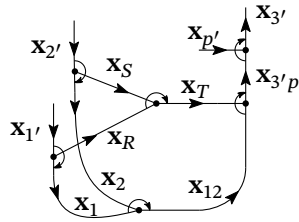
When acting with the effective Hamiltonian upon \mathbf{M} it is the goal to deform the symmetry tensors in such way that the resulting tensor \mathbf{H}_{part} has the same symmetry tensor structure as the original one. The contraction ends in the

same convention as it begins; it allows a straightforward comparison of the reduced tensors. For deforming the symmetry tensors we can do



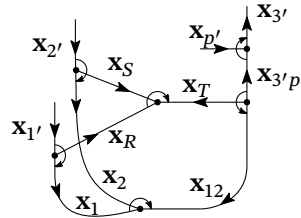
by using the orthogonality and reversal property

$$= (-1)^{-j_p+j_3-j_{12}} [j_{12}]^{-2} \delta_{x_{12}x_{3p}}$$



by using the reversal property twice

$$= (-1)^{-j_p+j_3+j_{12}+2j_T} [j_{12}]^{-2} \delta_{x_{12}x_{3p}}$$



And by finally using the 9j-merge property we obtain

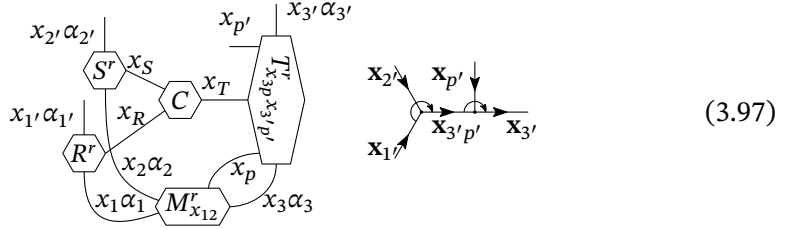
$$= (-1)^{-j_p+j_3+j_{12}+2j_T} [j_{12}]^{-2} \delta_{x_{12}x_{3p}} \left\{ \begin{matrix} j_R & j_S & j_T \\ j_1 & j_2 & j_{12} \\ j_{1'} & j_{2'} & j_{3'p'} \end{matrix} \right\} \begin{matrix} x_2 \searrow & x_{p'} \searrow \\ & \nearrow x_{3'p'} \\ x_{1'} \nearrow & & x_3 \end{matrix} .$$

The fermionic sign can be again calculated as in the previous section by writing out the states explicitly. Combining the fermionic sign and the SU(2)-prefactor

we have:

$$\mathbf{H}_{\text{part}} = \begin{array}{c} x_{2'}\alpha_{2'} \\ x_{1'}\alpha_{1'} \end{array} \begin{array}{c} x_p \\ x_{3'}\alpha_{3'} \end{array} \begin{array}{c} H_{x_3'p'}^r \\ \end{array} \begin{array}{c} x_{2'} \\ x_{1'} \end{array} \begin{array}{c} x_{p'} \\ x_{3'} \end{array} \begin{array}{c} x_{3'} \\ x_{3'} \end{array} \quad (3.96)$$

$$= (-1)^{\pi_p + \pi_R} \pi_2 (-1)^{-j_p + j_3 + j_{12} + 2j_T} [j_{12}]^{-2} \delta_{x_{12}x_{3p}} \begin{Bmatrix} j_R & j_S & j_T \\ j_1 & j_2 & j_{12} \\ j_{1'} & j_{2'} & j_{3'p'} \end{Bmatrix}$$



$$\quad (3.97)$$

The different contributions of the renormalized operators in Eq. (3.93) can be straightforwardly summed.

3.4.4. The eigenvalue problem

Although already discussed in the previous chapter, let us briefly revisit the eigenvalue problem for the optimization of a two-site tensor as there are some intricacies involved when using $SU(2)$ -adapted tensors.

In Section 3.2.2 we obtained the tools to build an $SU(2)$ -symmetric wave function in canonical form. In Section 3.3 we learned how the chemical Hamiltonian can be written in a $SU(2)$ -symmetric form and can be transformed to mimic the T3NS geometry; it allows an efficient evaluation of $\langle \Psi | H | \Psi \rangle$ which in turn provides an efficient minimization of the given Lagrangian

$$\mathcal{L} = \langle \Psi | H | \Psi \rangle - \lambda \langle \Psi | \Psi \rangle. \quad (3.98)$$

Let us revisit the optimization of the two-site tensor \mathbf{M} given in Eq. (3.84). When evaluating $\langle \Psi | \Psi \rangle$, we obtain

$$\langle \Psi | \Psi \rangle = \mathcal{C}(\mathbf{M}\mathbf{M}^+) \quad (3.99)$$

$$= (-1)^{\pi_3} (-1)^{j_1 + j_2 + j_p - j_3} \begin{array}{c} M_{x_{1'2'}}^r \\ x_{2'}\alpha_{2'} \\ x_p \\ M_{x_{12}}^r \\ x_{1'}\alpha_{1'} \end{array} \begin{array}{c} x_{3'}\alpha_{3'} \\ x_{1'} \\ x_{12} \\ x_{3'} \end{array} \quad (3.100)$$

$$= \sum_{x_1\alpha_1; x_2\alpha_2; x_{12}; x_p; x_3\alpha_3} [j_{12}]^{-2} \left| M_{x_1\alpha_1; x_2\alpha_2; x_{12}; x_p; x_3\alpha_3}^r \right|^2 \quad (3.101)$$

as \mathbf{M} is the current orthogonality center of the wave function. Note the extra phase $(-1)^{\pi_3}(-1)^{j_1+j_2+j_p-j_3}$ needed for the adjoint \mathbf{M}^+ . Just as the one-site orthogonality center in Eq. (3.49) this sets the adjoint \mathbf{M}^+ equal to the hermitian conjugate of \mathbf{M} up to an extra $(-1)^{\pi_3}$ phase due to the fermionic nature. While no rescaling was needed when calculating the norm of a one-site orthogonality center in Eq. (3.51), for a two-site tensor we need to rescale to take into account the internal edge.

For the evaluation of $\langle \Psi | H | \Psi \rangle$, we need to contract the tensor \mathbf{M} with different sets of three renormalized operators as discussed in the previous section; \mathbf{H} is the result of the action of this effective Hamiltonian on \mathbf{M} . The result is given in Eq. (3.97). Note that the structures of the symmetry tensor of \mathbf{M} in Eq. (3.84) and \mathbf{H} are the same; it allows a one-to-one comparison of the reduced elements and with minimal effort one can show

$$\langle \Psi | H | \Psi \rangle = \mathcal{C}(\mathbf{H}\mathbf{M}^+) = \sum_{x_1\alpha_1;x_2\alpha_2;x_{12};x_p;x_3\alpha_3} [j_{12}]^{-2} \bar{M}_{x_1\alpha_1;x_2\alpha_2;x_{12};x_p;x_3\alpha_3}^r H_{x_1\alpha_1;x_2\alpha_2;x_{12};x_p;x_3\alpha_3}^r \quad (3.102)$$

The results in Eq. (3.101) and Eq. (3.102) can be used to minimize the Lagrangian with respect to the tensor \mathbf{M} ; it reduces the problem to solving an ordinary eigenvalue problem. However, one should take into account the scale-factor $[j_{12}]^{-1}$ before feeding it to a sparse eigensolver; i.e. for any two-site tensor \mathbf{A}

$$A^r \xrightarrow{\text{to solver}} [j_{12}]^{-1} A^r \quad (3.103)$$

$$A^r \xrightarrow{\text{from solver}} [j_{12}] A^r \quad (3.104)$$

where we need to scale either up or down for results we extract from or feed to the eigensolver. This is inherent to absorbing the $[j_i]$ as discussed on page 61. When explicitly keeping track of these factors, a rescaling of the tensors is not needed. One cannot completely forget about the $[j_i]$ terms either way.

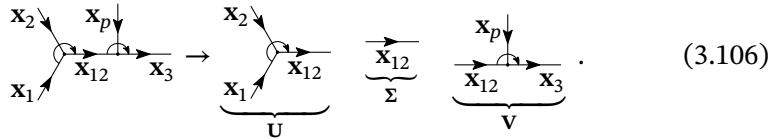
3.4.5. Singular value decomposition with $SU(2)$ -symmetry

In the previous sections and in Section C.3 of Appendix C an almost complete toolbox has been described for the implementation of an $SU(2)$ -symmetric T3NS algorithm. The only missing piece is a reformulation of the singular value decomposition such that we can split the optimized multi-site tensor in a controlled way; it allows a dynamic selection of important renormalized states on an entanglement based approach in symmetry-adapted networks. To illustrate SVD in this $SU(2)$ -adapted approach, we will again look at the multi-site tensor \mathbf{M} given in Eq. (3.84). A decomposition

$$\mathbf{M} = \mathcal{C}(\mathbf{U}\Sigma\mathbf{V}) \quad (3.105)$$

is obtained where \mathbf{U} and \mathbf{V} are suitable orthogonalized tensors and Σ are the singular values. First, we assign the outer indices of \mathbf{M} to either \mathbf{U} or \mathbf{V} , specifying the decomposition. The decomposition becomes particularly straightforward when the SVD partitioning corresponds with cutting one internal edge of the symmetry tensors of \mathbf{M} . If this is not the case, we can make it so by first deforming the symmetry tensors using their properties; hence, we will only discuss the former case.

Cutting along the internal degree of freedom x_{12} of \mathbf{M} , We can easily split the symmetry tensors into a part for \mathbf{U} and \mathbf{V} and a part for the singular values, i.e.



$$\begin{array}{c} x_2 \swarrow \\ \text{---} \text{---} \\ \nearrow x_1 \end{array} \begin{array}{c} x_p \downarrow \\ \text{---} \text{---} \\ \nearrow x_{12} \end{array} \begin{array}{c} \text{---} \text{---} \\ \text{---} \text{---} \\ \searrow x_3 \end{array} \rightarrow \underbrace{\begin{array}{c} x_2 \swarrow \\ \text{---} \text{---} \\ \nearrow x_1 \end{array}}_{\mathbf{U}} \underbrace{\begin{array}{c} x_{12} \leftarrow \\ \text{---} \text{---} \\ \rightarrow x_{12} \end{array}}_{\Sigma} \underbrace{\begin{array}{c} x_p \downarrow \\ \text{---} \text{---} \\ \searrow x_3 \end{array}}_{\mathbf{V}} \quad (3.106)$$

Now the values of the reduced tensors can be calculated through an ordinary SVD; remember that U^r and V^r should be orthogonal up to multiplicities as seen in Section 3.2.2; we get

$$M_{x_1\alpha_1;x_2\alpha_2;x_{12}\alpha_{12};x_p;x_3\alpha_3}^r = \sum_{\alpha_m} [j_{12}] U'_{x_1\alpha_1;x_2\alpha_2;x_{12}\alpha_{12}} [j_{12}]^{-2} \Sigma'_{x_{12}\alpha_{12}} [j_{12}] V'_{x_{12}\alpha_{12};x_p;x_3\alpha_3} \quad (3.107)$$

Or thus

$$U_{x_1\alpha_1;x_2\alpha_2;x_{12}\alpha_{12}}^r = [j_{12}] U'_{x_1\alpha_1;x_2\alpha_2;x_{12}\alpha_{12}} \quad (3.108)$$

$$\Sigma_{x_{12}\alpha_{12}}^r = [j_{12}]^{-2} \Sigma'_{x_{12}\alpha_{12}} \quad (3.109)$$

$$V_{x_{12}\alpha_{12};x_p;x_3\alpha_3}^r = [j_{12}] V'_{x_{12}\alpha_{12};x_p;x_3\alpha_3} \quad (3.110)$$

The presented decomposition of the symmetry tensors in Eq. (3.106) and the above defined reduced tensors provide a valid singular value decomposition; the thorough reader can easily check that Eq. (3.105) is fulfilled and both \mathbf{U} and \mathbf{V} are suitably canonicalized (orthogonal) tensors.

Last, it is worth pointing out one important caveat. The singular values in $\Sigma_{x_{12}\alpha_{12}}^r$ correspond with the singular values of the irreducible tensor \mathbf{M} ; however each value of $\Sigma_{x_{12}\alpha_{12}}^r$ has an implied multiplicity of $(2j_{12} + 1)$. When truncating the two-site tensor up to a certain number of reduced states one should take this multiplicity into account; some small singular values can have larger contributions to the two-site tensor due to their higher multiplicity than other.

3.5. Numerical results

In this section, we present several calculations with T3NS. All tensors are randomly initialized. The used implementation for the calculations can be found

on github [156]. This implementation is able to exploit \mathbb{Z}_2 , $U(1)$, $SU(2)$ and the real abelian point group symmetries. The symmetries can be included in a modular way. This enables us to compare calculations with and without $SU(2)$ and point group symmetries for the same implementation. The presented calculations were previously published in Ref. [105].

After the optimization of two sites, this two-site tensor has to be split into two separate sites again. This is done by SVD. The truncation of the virtual bond dimension during this step is done in two different ways.

First, a fixed maximal bond dimension can be imposed. The algorithm will take as many singular values into consideration as possible. It will first select the largest singular values until no non-zero singular values are left or the maximal bond dimension is reached. The remaining singular values and their corresponding basis states are discarded.

Second, the dynamic block state selection (DBSS) can be used [149, 192]. With this method, the algorithm keeps the largest singular values until a certain threshold for a cost function is reached. The cost function used in our implementation is given by

$$w_{\text{disc}} = \sum_{i_{\text{disc}}} s_{i_{\text{disc}}}^2 \quad (3.111)$$

i.e. the sum of the squares of all discarded singular values. This corresponds with $\langle \Psi_{\text{disc}} | \Psi_{\text{disc}} \rangle$, with $|\Psi_{\text{disc}}\rangle$ the discarded part of the wave function during truncation. Other cost functions can be easily implemented. Next to a targeted threshold, a minimal and maximal bond dimension should be specified. Throwing away too many basis states at a certain stage can impede the optimization at later stages, even though the truncation error is only minimal at that point. Specifying a minimal bond dimension ensures a certain flexibility at all time. The maximal bond dimension is needed to prevent a large increase in both run time and memory usage when the imposed threshold can not be reached.

As noted in Section 3.1.2.3, the usage of symmetries with irreps that are more than one-dimensional, such as $SU(2)$, introduces a compression of the wave function. Different basis states belonging to the same multiplet can be represented by a singular reduced basis state. Analogous to the bond dimension being the number of basis states kept in the bond, the *reduced bond dimension* is defined as the number of reduced basis states kept. When using $SU(2)$ the reduced bond dimension of the bonds, and not the bond dimension, will reflect the computational complexity.

Just as with DMRG for quantum chemistry, keeping track of the renormalized operators is the most taxing part on memory for T3NS. The amount of

renormalized operators needed to be stored for T3NS are of the same order as for DMRG. However, T3NS calculations can be performed on a considerably lower bond dimension for a similar accuracy. Consequently, this lowers the storage requirements for the renormalized operators and allows us to keep all tensors on memory at all time. No checkpoint files need to be written to disk or read from disk during the algorithm for the present system sizes and bond dimensions.

3

3.5.1. The Bisoxo and Peroxo Isomer of $[\text{Cu}_2\text{O}_2]^{2-}$

We revisit the bis(μ -oxo) and the $\mu = \eta^2 : \eta^2$ peroxo $[\text{Cu}_2\text{O}_2]^{2-}$ isomers as a test case for T3NS with SU(2) and abelian point group symmetries. These transition metal complexes have been previously studied by other ab initio methods such as the complete active space self-consistent field theory (CASSCF), the complete active space self-consistent field theory with second order perturbation theory (CASPT2) [194], the restricted active space self-consistent field theory with second order perturbation theory (RASPT2) [195], DMRG [142, 149] and DMRG+CT [150] (DMRG with canonical transformation theory). It was the largest system studied in the initial T3NS paper using only $U(1) \otimes U(1)$ symmetry [104] (i.e. conservation of both particle number and spin projection).

We perform calculations for both isomers in an (26e, 44o) active space. The basis set is given by the Stuttgart pseudopotential and associated basis functions (ECP10MDF) [201] for Cu and the atomic natural orbital (ANO) basis set of Pierloot *et al.* for O [202]. The same active space is used as in Refs. [104, 149]. Both isomers have a D_{2h} point group symmetry and their ground state is a singlet state in the A_g irrep of D_{2h} [149]. When using the SU(2) and/or point group symmetry adapted version of T3NS, states corresponding to these irreps will be targeted. The same T3NS shape and orbital ordering is used as in Ref. [104]. We also perform calculations of the lowest lying triplet state in the A_g irrep of D_{2h} for both isomers.

The Bisoxo isomer with and without spin symmetry

In order to compare the present spin adapted version of T3NS with its non-adapted predecessor [104], we perform several calculations for the bisoxo isomer at different bond dimensions and with different symmetries included.

In Fig. 3.2a, timings for the last sweep are shown for several fixed bond dimensions. Calculations were performed with $U(1) \otimes U(1)$ symmetry and with $SU(2) \otimes U(1)$ symmetry. For both, calculations with and without the D_{2h} point group symmetry are done. For the spin-adapted versions, the bond dimensions shown are the reduced bond dimensions.

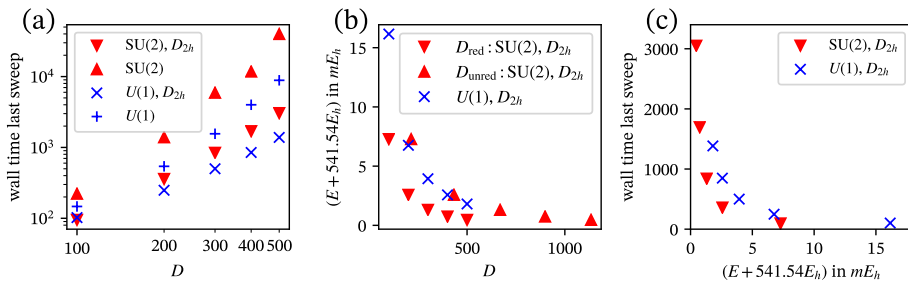


Figure 3.2: Results for the bisoxo isomer obtained on an 18-core Intel Xeon Gold 6140 (Skylake at 2.3 GHz). Calculations with a fixed maximal bond dimension of $D = 100, 200, 300, 400$ and 500 were performed. Results for $U(1) \otimes U(1)$ combined with or without D_{2h} are given by crosses. Results for $U(1) \otimes SU(2)$ combined with or without D_{2h} are given by triangles. The bond dimensions are in this case the reduced bond dimensions, except for (b) where also unreduced bond dimensions are shown. (a) shows the wall time of the last sweep for different bond dimensions and different symmetries. (b) shows the ground state energy for different bond dimensions. For the spin-adapted calculations, both the maximal reduced as unreduced bond dimensions are given. (c) shows wall time of last sweep in function of ground state energy.

As expected, the usage of the point group symmetry introduces a lot of sparsity in the tensors which speeds up calculations considerably. Calculation time improved by a factor of 6 and 13 at $D = 500$ for $U(1) \otimes U(1)$ and $SU(2) \otimes U(1)$ respectively when including D_{2h} . For both calculations with or without spin symmetry, the inclusion of D_{2h} yields practically the same energies and maximal truncation errors as when performing the calculation without the point group symmetry.

For tensors of the same size, calculations including spin symmetry are computationally more intensive than without spin symmetry as can be seen in Fig. 3.2a. This is expected since the reduced tensors are more dense than the irreducible tensors. However, the compressed nature of the reduced tensors (see section 3.1.2.3) makes spin-adapted calculations at a certain *reduced* bond dimension more accurate than calculations without spin symmetry at an equal bond dimension. This can be seen in Fig. 3.2b. The maximal unreduced bond dimension during $SU(2)$ calculations is also given in this figure. For the present calculations, the maximal unreduced bond dimension is approximately twice as large as the imposed maximal reduced bond dimension. When comparing wall time with achieved accuracy, the calculations with $SU(2)$ are considerably faster, as is shown in Fig. 3.2c.

The bisoxo and peroxo isomers with spin symmetry

Several calculations are performed for both isomers. Both $SU(2)$ and D_{2h} symmetry are used. The inclusion of spin and point group symmetry considerably improves our calculations and allows us to go to much larger bond dimensions. Both a constant maximal bond dimension and DBSS are used. Some obtained results are shown in Table 3.2 alongside previously published results. Calculations at a fixed reduced bond dimension of 300 already surpassed the most accurate DMRG calculations of Ref. [149] and the most accurate T3NS calculations of Ref. [104]. For the most accurate calculation the maximal reduced bond dimensions needed were 1626 and 1329 for the bisoxo and peroxo isomer respectively.

Method	$E_{\text{bisoxo}}[E_h]$	$E_{\text{peroxo}}[E_h]$	ΔE [kcal/mol]
DMRG[149]	-541.538 53	-541.581 14	26.7
T3NS with $U(1) \otimes U(1)$ [104]			
500	-541.538 20	-541.580 94	26.8
T3NS with $U(1) \otimes SU(2) \otimes D_{2h}$			
300	-541.538 69	-541.581 19	26.7
500	-541.539 54	-541.581 71	26.5
500, 5×10^{-6}	-541.539 86	-541.581 83	26.3
1000	-541.539 93	-541.581 97	26.4
1000, 5×10^{-6}	-541.539 97	-541.581 98	26.4
Extrapolated	-541.540 12	-541.582 10	26.3

Table 3.2: Energy gaps and ground state energies between the bisoxo and peroxo isomers. The energy gaps are given in kcal/mol. Ground state energies are given in Hartree. For T3NS calculations with fixed bond dimension, the bond dimension is given in the first column. For T3NS calculations using DBSS, the minimal bond dimension and the truncation error is given. For T3NS with $SU(2)$ the given bond dimensions are the reduced ones. Maximum bond dimensions around 2000 were reported for both clusters for the DMRG calculations of Ref. [149].

A linear extrapolation [182, 204] between the truncation error and the energy is performed. The extrapolation is based on results with a fixed maximal reduced bond dimension of $D = 600, 700, 800, 900$ and 1000. This extrapolation is shown in Fig. 3.3. The results for the DBSS calculations in Table 3.2 are also given in the figure. For both systems, two DBSS calculations are performed, both targeting a truncation error of 5×10^{-6} . They use however another minimal bond dimension. Although the same truncation error is targeted for both DBSS calculations, the accuracy of the two calculations is quite different due to their different minimal bond dimensions. Because of this, only the calculations with fixed bond dimension are used for the extrapolation. A

linear extrapolation seems justified for these calculations. The extrapolated values are given in Table 3.2.

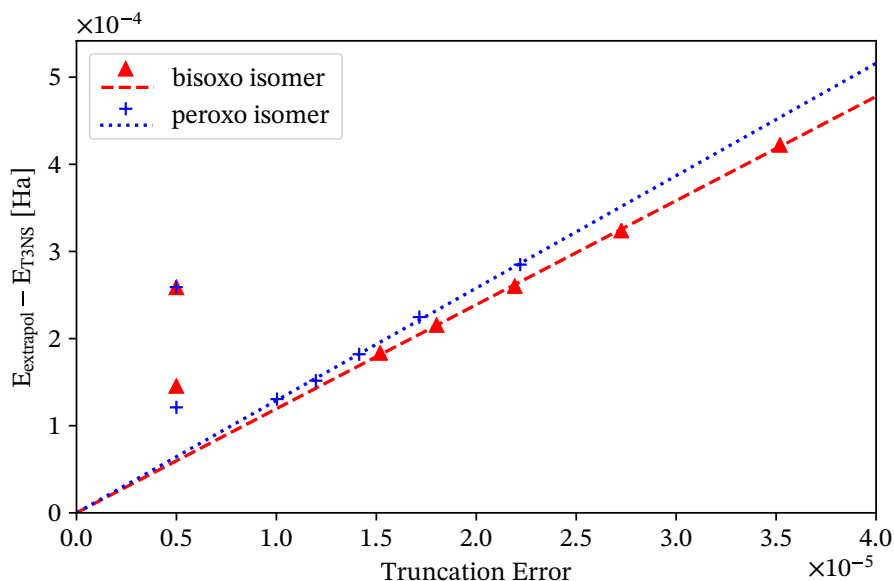


Figure 3.3: Extrapolation of the energy for both the bisoxo and peroxo isomer. Extrapolation is done by using the results for $D = 600, 700, 800, 900$ and 1000 . The DBSS calculations of Table 3.2 targeting a truncation error of 5×10^{-6} are also shown in the figure. They are however not used for the extrapolation.

The lowest lying triplet states in the A_g irrep are also calculated for both isomers. These states are easily targeted by changing the allowed quantum numbers in the outgoing target state of the T3NS (as discussed in Section 3.2.3) and are similar in computation time as for the singlet states. We obtained an energy of $E = -541.46194 E_h$ and $E = -541.40184 E_h$ for the bisoxo and peroxo isomer respectively when a maximal reduced bond dimension of 1000 was chosen.

3.6. A short recap

In this chapter, we studied how the $SU(2)$ -symmetry of the chemical Hamiltonian can be exploited in the T3NS through the Wigner-Eckart theorem. By expressing all renormalized and physical states as angular momentum eigenstates the tensors can be decomposed into a reduced tensor and a symmetry tensor. As the latter is dictated by the Clebsch-Gordan coefficients of the symmetries the optimization is restricted to the reduced tensor elements which are

independent of the exact spin projection of the renormalized states. Not only the number of parameters diminishes by exploiting the $SU(2)$ -symmetry but contractions can also be done more efficiently by leveraging useful properties of the Clebsch-Gordan coefficients. While the numerical benefits are already substantial, exploiting spin symmetries also enable the targeting of specific spin states. The lowest lying eigenstate for each total spin is obtained through a straightforward ground state calculation.

3

Not only the $SU(2)$ -symmetry can be exploited, but we also make usage of any real abelian point group symmetries present in the molecule. While the implementation of these symmetries is much simpler, their usage can result in vast speedups when present. The chapter was concluded by some numerical illustrations.

As a final remark, we would like to reiterate that the used convention for the implementation of the $SU(2)$ -symmetry in this chapter defers to the one used in the source code [156]. The exact differences and the motivation for this change of convention are explained on page 61. In short, the convention used in this dissertation has fewer terms to explicitly keep track of, resulting in it being less error-prone and more generalizable to work with.

4

Seniority

*Two left-handed gloves don't make a pair.
Two half-truths don't make a truth.*

Multatuli

In the previous chapter we studied how symmetries of the Hamiltonian can be leveraged for a faster and more accurate T3NS algorithm. In this chapter we will apply the same techniques to the seniority of the wave function. Defined as the number of unpaired electrons, the seniority is not a symmetry of the chemical Hamiltonian. However, by writing the different tensors in the tensor network invariant for the seniority we get a natural way of restricting the ground state optimization to seniority subspaces. It allows a polynomial scaling method for the study of molecular systems while imposing a restriction on the seniority of their wave function.

Tensor networks also provide an efficient way of finding the ground state within the doubly occupied configuration interaction (DOCI) subspace. This is particularly useful as it has been shown that most of the strong correlations present in a molecule are already present in this subspace.

This chapter is based on a paper published as a preprint on arXiv [221].

The concept of seniority – originally studied in nuclear and condensed matter physics [222, 223] – has proven to be quite useful for capturing strong correlations in quantum chemical systems. The seniority of a wave function is defined as the number of unpaired electrons; we call two electrons paired when occupying the same spatial orbital. E.g., the following wave function in a space of two spatial orbitals has a seniority $\nu = 0$:

$$|\Psi\rangle = |\uparrow\downarrow, -\rangle = c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger |-\rangle. \quad (4.1)$$

It is an example of a doubly occupied configuration interaction (DOCI) wave function as the spatial orbitals are either doubly occupied or completely empty. For molecular systems, it was shown that a DOCI wave function often suffices to capture most of the strong correlation present [224–230]. Although this tremendously reduces the dimension of the Hilbert space, finding the exact DOCI wave function is still exponentially scaling as assigning $N/2$ electron pairs to L spatial orbitals is a combinatorial problem. Furthermore, the seniority of a wave function depends on the choice of orbitals; seniority subspaces are not invariant under orbital rotations as illustrated by the following rotation for Eq. (4.1):

$$c_{1\sigma}^\dagger = \frac{1}{\sqrt{2}}(\tilde{c}_{1\sigma}^\dagger + \tilde{c}_{2\sigma}^\dagger), \quad c_{2\sigma}^\dagger = \frac{1}{\sqrt{2}}(-\tilde{c}_{1\sigma}^\dagger + \tilde{c}_{2\sigma}^\dagger). \quad (4.2)$$

This results into

$$|\Psi\rangle = c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger |-\rangle = \frac{1}{2}[\tilde{c}_{1\uparrow}^\dagger \tilde{c}_{1\downarrow}^\dagger + \tilde{c}_{1\uparrow}^\dagger \tilde{c}_{2\downarrow}^\dagger + \tilde{c}_{2\uparrow}^\dagger \tilde{c}_{1\downarrow}^\dagger + \tilde{c}_{2\uparrow}^\dagger \tilde{c}_{2\downarrow}^\dagger] |-\rangle \quad (4.3)$$

$$= \frac{1}{2}[|\uparrow\downarrow, -\rangle + |\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle + |-, \uparrow\downarrow\rangle] \quad (4.4)$$

which is a mixture of Slater determinants with seniority $\nu = 0$ and $\nu = 2$.

At first glance, DOCI seems only marginally more manageable than exactly solving the Schrödinger equation as DOCI is both exponential scaling and orbital dependent. The antisymmetric product of one-reference orbital geminals (AP1roG) [228, 231–233] – also known as pair coupled cluster doubles (pCCD) [234–236] – provides a reliable approximation to the DOCI ground state solution for a wide range of molecular systems while staying computationally tractable. Soon after, the first orbital optimization schemes [232, 233] in accompaniment with AP1roG were formulated; an indispensable tool for optimizing the orbital dependent seniority-zero subspace. However, even with the flexibility of orbital optimization, solutions living within the seniority-zero subspace lack some nondynamical strong correlations (as can be seen in the underestimated binding energy of the nitrogen dimer) and dynamical weak

correlations. Although several extensions to AP1roG exist trying to recover the missing weak correlations, they miss the same reliability as AP1roG has for the strong correlation [237].

Not only the seniority-zero subspace, but also higher seniority subspaces have been studied in a molecular setting, albeit only in small systems [226, 229, 234, 238, 239]. The seniority number was also used for the definition of a renormalization flow within the energy renormalization group [240]. However, these methods are exponentially scaling if there is no truncation in the Slater determinants considered.

In this chapter we will use the T3NS to study the behavior of chemical systems in a seniority-restricted subspace. The T3NS considers the full basis of Slater determinants during its optimization, instead it truncates the amount of entanglement captured. This efficient many-body technique will enable us to optimize wave functions for large systems in seniority-restricted subspaces and is not plagued by the exponential wall as previous seniority-restricted studies [226, 229, 234, 238–240]. In the next section we will discuss how to leverage the techniques discussed in the previous chapters to perform seniority-restricted calculations.

In Section 4.2.1 we study the dissociation of the nitrogen dimer. While decently described in DOCI, we argue and show why at least seniority-four has to be used for a quantitative correct dissociation. In previous research, orbital optimized DOCI has introduced an artificial point group symmetry breaking for the benzene molecule; in Section 4.2.2 we study at which seniority sector this artifact is resolved. In Section 4.2.3 we discuss the neon dimer, a dimer merely binding due to weak dispersion forces. As DOCI is not fit for dynamical correlations, we study which seniority subspaces need to be included for the retrieval of the right dissociation character. We conclude this chapter by showing some large-scale DOCI-T3NS calculations in Section 4.2.4. Due to the increased simplicity of the DOCI-restricted Hamiltonian, we can push these types of calculations to much larger molecules than possible for general quantum chemical TNS calculations.

4.1. Seniority and tensor networks

In Chapter 3 we investigated how T3NS can exploit the symmetries showcased by the chemical Hamiltonian. In this chapter, we will apply the same techniques to allow seniority-restricted calculations. We write each tensor in the network in an invariant form for the seniority which behaves the same as the particle conservation, i.e. it is an $U(1)$ -symmetry.

As a refresher, let us reiterate how the invariant form can be obtained for a three-legged tensor. We will solely concentrate on the seniority in the discus-

sion to follow; bear in mind that we can easily use seniority in junction with symmetries of the Hamiltonian such as spin conservation. We need to impose the following restriction on the tensor elements:

$$\begin{array}{c} \downarrow b \\ \circlearrowleft (A) \\ \leftarrow a \quad \rightarrow c \end{array} = 0, \quad \text{if } \nu_a + \nu_b \neq \nu_c, \quad (4.5)$$

where ν_a , ν_b and ν_c are the seniority numbers of their respective (physical or virtual) degrees of freedom. The seniority number of the first two degrees of freedom a and b have to sum up to the seniority number of c . This restriction implies a kind of *flow* for the seniority number in the network as indicated in Eq. (4.5). This is of course the same flow as encountered in the previous chapter. An example of an MPS built from three of these invariant tensors with the flow indicated is given by:

$$\begin{array}{c} \downarrow a \\ \circlearrowleft (A) \\ \leftarrow \text{vac} \quad \rightarrow \alpha \end{array} \begin{array}{c} \downarrow b \\ \circlearrowleft (B) \\ \leftarrow \alpha \quad \rightarrow \beta \end{array} \begin{array}{c} \downarrow c \\ \circlearrowleft (C) \\ \leftarrow \beta \quad \rightarrow \text{target} \end{array}. \quad (4.6)$$

Here the physical edges are denoted with a , b and c which have a seniority of $\nu \in \{0, 1\}$ as each orbital can be either empty ($\nu = 0$), singly occupied ($\nu = 1$) or double occupied ($\nu = 0$). The vacuum state entering on the leftmost side of the network is characterized by a seniority $\nu_{\text{vac}} = 0$. Up to now, there is no difference in how one would implement a real $U(1)$ -symmetry of the system such as a particle conservation or conservation of the spin projection. However, the seniority is *not* a conserved quantum number. Eigenstates of the Hamiltonian are not necessarily eigenstates of the seniority operator and the target state will generally be a linear combination of Slater determinants with different seniority numbers. To target such a state, the final renormalized states in the rightmost bond should be a set of eigenstates of the seniority operator which combine to the targeted state. The set of possible seniority numbers for the wave function is

$$\Omega = \left\{ n \in \mathbb{N} : \begin{array}{l} n \bmod 2 = N_{\text{tot}} \bmod 2 \\ |N_{\uparrow} - N_{\downarrow}| \leq n \leq \min(N_{\text{tot}}, 2L - N_{\text{tot}}) \end{array} \right\}, \quad (4.7)$$

with L the number of spatial orbitals, N_{\uparrow} (N_{\downarrow}) the number of electrons with spin up (down) and N_{tot} the total number of electrons. For every renormalized state in the last bond, we have $\nu_{\text{target}} \in \Omega$. By restricting ν_{target} to a subset $S \subseteq \Omega$, ground states in seniority-restricted subspaces can be targeted.

In a similar fashion, one could also use other non-conserved quantum numbers than the seniority number. For example, we could use the excitation number with respect to the Hartree-Fock wave function; by only allowing

Slater determinants with a certain amount of excitations, tensor networks can be used as an approximate configuration interaction (CI) solver with arbitrary allowed excitation levels.

4.1.1. Suboptimal decomposition

When using a wave function ansatz as shown in Eq. (4.6), we impose a restriction on the left renormalized states at each splitting of the network. Due to the fact that a vacuum state enters in the leftmost bond and all used tensors are invariant under the seniority operator, the left renormalized states need to have a well defined seniority number. This restriction results in the need for a possibly larger bond dimension than when discarding seniority. We illustrate this using a wave function with three electrons and three spatial orbitals:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} [|\uparrow, \downarrow\rangle \otimes |\uparrow\rangle + |-\, \uparrow\downarrow\rangle \otimes |\uparrow\rangle] \quad (4.8)$$

$$= \frac{1}{\sqrt{2}} [|\uparrow, \downarrow\rangle + |-\, \uparrow\downarrow\rangle] \otimes |\uparrow\rangle. \quad (4.9)$$

In Eq. (4.9), the Schmidt decomposition for a partitioning between the first two and the last orbital is given. At this partitioning only a virtual bond dimension of one is needed to represent the state. However, when we impose that the left states, i.e. the states in the first two orbitals, of the decomposition should also be eigenstates of the seniority operator, the needed bond dimension at this partitioning increases to two, cf. Eq. (4.8).

In the previous example, the sink (i.e. the outgoing virtual bond of the tensor network) is positioned at the rightmost bond of the tensor network cf. Eq. (4.6). One could also opt to position the sink at another point in the tensor network. Graphically this can be done by using the following kind of T3NS:

$$\begin{array}{ccccccc} & & & \text{target} & & & \\ & & & \uparrow & & & \\ \text{vac} & \xrightarrow{a} & \textcircled{A} & \xrightarrow{\alpha} & \textcircled{B} & \xrightarrow{\beta} & \textcircled{E} & \xrightarrow{\delta} & \textcircled{D} & \xrightarrow{\gamma} & \textcircled{C} & \xrightarrow{\text{vac}} \\ & & & & & & & & & & & \end{array} \quad (4.10)$$

Here, two MPS-chains are connected through a branching tensor E in the middle. The sink is now positioned at this branching tensor.

We numerically illustrate the effect of using seniority-restricted tensors on the von Neumann entropy – which can be seen as a measure for the needed bond dimension – by solving the LiF molecule. The same basis set is used as in Section 2.8 and all calculations can be considered exact. In Fig. 4.1, results for DMRG calculations with seniority-invariant and non-invariant tensors are shown. The effect on the entropy of the position of the sink is also illustrated

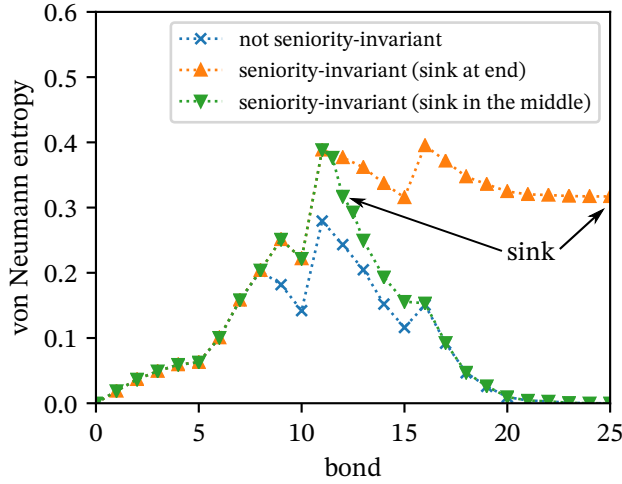


Figure 4.1: The von Neumann entropy for LiF along each virtual bond in an MPS calculations is compared with and without seniority-invariant tensors. The same basis set is used as in Section 2.8. The bonds are given in order of the MPS, i.e. bond i is the i 'th bond starting from the left of the MPS. All MPS calculations use the same orbital ordering.

by comparing calculations with the sink at the end (cf. Eq. (4.6)) and with the sink in the middle of the tensor network (cf. Eq. (4.10)). As expected the von Neumann entropy is larger when imposing seniority-invariance; furthermore, when the sink is positioned at the end of the chain we notice the worst accumulation of entanglement due to the artificial separation of renormalized states with different seniority. The entanglement at the sink is a fixed value irrespective of the location in the network and is given by $S = \sum_{\nu} -|c_{\nu}|^2 \log|c_{\nu}|^2$ where $|c_{\nu}|^2$ are the weights of the different seniority subspaces in the wave function.

4.1.2. DOCI and tensor networks

Restricting the calculation to configurations with $\nu = 0$ is easily done with the aforementioned method. However, it is more efficient to directly implement the quantum chemical Hamiltonian projected on the DOCI-subspace where only paired electrons are allowed. The DOCI-Hamiltonian is given by

$$H_{\text{DOCI}} = 2 \sum_i T_{ii} n_i + \sum_{ij} [2V_{ijij} - V_{ijji}] n_i n_j + \sum_{i \neq j} V_{iijj} b_i^{\dagger} b_j, \quad (4.11)$$

where b_i^\dagger and b_i are the bosonic pair creation and annihilation operators and n_i is the pair number operator at orbital i . They are given by

$$b_i^\dagger = c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger, \quad b_i = c_{i\downarrow} c_{i\uparrow} \quad (4.12)$$

and

$$n_i = b_i^\dagger b_i. \quad (4.13)$$

This Hamiltonian only scales quadratic with the number of orbitals in contrast with the quartic scaling of the full Hamiltonian; as such, TNS calculations in the DOCI subspace can be performed with a lower polynomial scaling. The scaling is given in Table 4.1 for both DMRG and T3NS.

	CPU time	Memory
QC-DMRG:	$\mathcal{O}(L^4 D^2 + L^3 D^3)$	$\mathcal{O}(L^3 D^2)$
DOCI-DMRG:	$\mathcal{O}(L^2 D^3)$	$\mathcal{O}(L^2 D^2)$
QC-T3NS:	$\mathcal{O}(L^4 D^2 + L^3 D^4)$	$\mathcal{O}(L^3 D^2 + L D^3)$
DOCI-T3NS:	$\mathcal{O}(L^2 D^4)$	$\mathcal{O}(L^2 D^2 + L D^3)$

Table 4.1: Resource requirements for DMRG and T3NS with renormalized operators for the full quantum chemical Hamiltonian in Eq. (1.18) or the DOCI Hamiltonian in Eq. (4.11) for L spatial orbitals. The maximal virtual bond dimension is denoted by D .

We find that DOCI ground state wave functions have in general lower entanglement than their corresponding FCI ground state wave function; accurate results for DOCI can be obtained with a lower bond dimension. The synergy between the lower polynomial scaling and the lower required bond dimension makes DOCI-TNS very fast. It makes DOCI-TNS a good option for initializing tensor network calculations in the FCI space; the DOCI-TNS optimization takes only a few minutes for systems considered large for QC-TNS. As an example, we will show some DOCI-T3NS calculations on a rubidium phenanthrene coordination complex in Section 4.2.4 in a minimal basis set. This system will provide a decent challenge for DOCI-TNS as it involves the optimization of 162 electron pairs in 261 spatial orbitals.

4.2. Applications

In this section, some calculations with the seniority-restricted tensor network code are discussed. As these calculations are orbital dependent, several types of orbitals are considered. The effect of allowing progressively more broken pairs is also studied within each orbital set. In Section 4.2.1 and Section 4.2.3, the dissociation of the nitrogen and neon dimer are considered, respectively.

Section 4.2.2 discusses the benzene molecule, a system demonstrating artificial D_{6h} symmetry breaking in the seniority-zero subspace [233]. In Section 4.2.4, we study a rubidium phenanthrene coordination complex in a minimal basis set with all electrons correlated. For this we only use the DOCI-TNS method as to show its speed with respect to full TNS calculations in quantum chemistry.

The coupled cluster natural orbitals and Löwdin orthogonalized atomic orbitals are obtained with PySCF [241–243]. The DOCI-optimized orbitals are generated by an in-house DOSCF code and were carefully checked to correspond to the lowest possible DOCI energy, i.e. the global minimum [244]. All seniority-restricted tensor network calculations are MPS calculations. We exploit the spin symmetry and the reported bond dimensions for the tensor networks are *reduced* bond dimensions; renormalized states belonging to the same multiplet are represented by one reduced renormalized state, thus reducing the needed bond dimension. Seniority-restricted tensor network calculations are, just as regular tensor network calculations, not exact; the accuracy can be controlled by the bond dimension. The following calculations use a large enough bond dimension such that results are within a high enough accuracy for quantitatively accurate potential energy surfaces.

4

4.2.1. Nitrogen dimer

Characterized by a triple bond breaking, the nitrogen dimer is a much visited test case for new quantum chemical methods. Here, we study the nitrogen dimer in a cc-pVDZ basis set with all electrons correlated. Seniority-restricted spin-adapted DMRG with a reduced bond dimension up to 1000 is used to optimize the ground state in the different subspaces. The allowed seniority increases from 0 (DOCI) up to 10 for the largest calculations, allowing 5 electron pairs to be broken. In Fig. 4.2, the dissociation curves are given for calculations within the different seniority subspaces. Calculations were performed for canonical orbitals (Fig. 4.2a), DOCI-optimized orbitals (Fig. 4.2b) and CCSD natural orbitals (Fig. 4.2c). Although the DOCI-optimized orbitals are optimized for the seniority-zero subspace specifically, they also perform better in higher seniority subspaces, albeit marginally. Eventually for $\nu \leq 8$ and onward, all orbital sets give quasi-FCI energies.

In Ref. [232], it is shown that, up to first order, the seniority-two sector decouples from the seniority-two-plus-zero sector for DOCI-optimized orbitals; only a small correction should occur when introducing single broken pairs in this orbital set. We indeed notice a small energy correction for the DOCI-optimized orbitals, smaller than for canonical orbitals. In Fig. 4.3, the weights of the different seniority subspaces are plotted for the ground state in both canonical and DOCI-optimized orbitals. It is yet another illustration that for

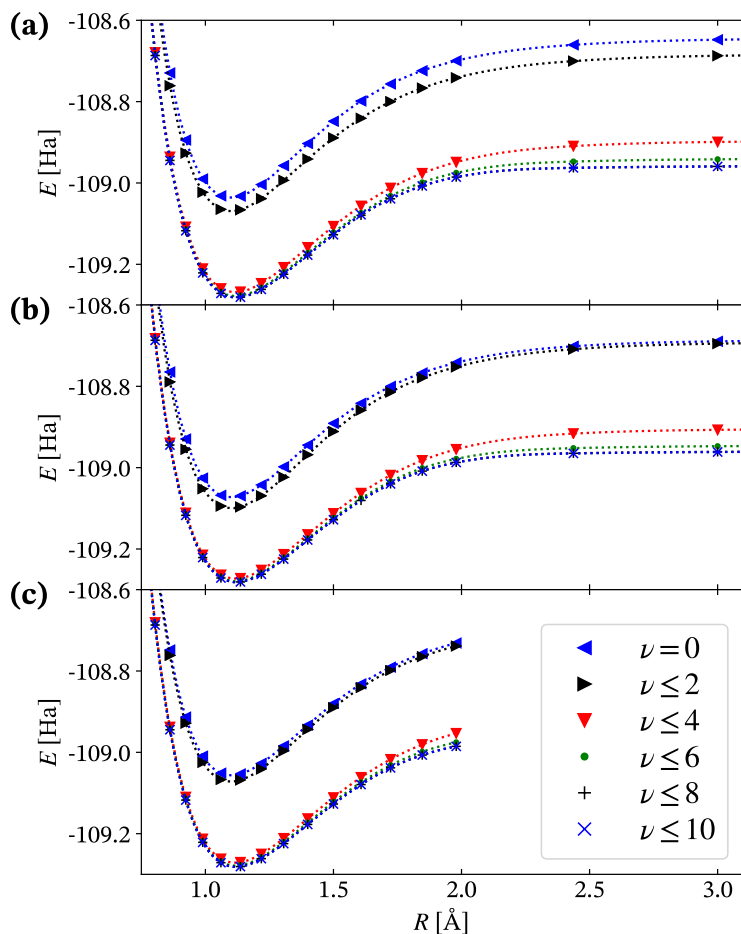


Figure 4.2: Dissociation curves at different seniority subspaces for the nitrogen dimer. Results for canonical orbitals (a), DOCI-optimized orbitals (b) and CCSD natural orbitals (c) are given. For (c), only results where CCSD converged are plotted.

DOCI-optimized orbitals (Fig. 4.3b) the seniority-two subspace is less important than for canonical orbitals (Fig. 4.3a). However, a first order decoupling is not an exact one; there are other orbital sets possible which give even smaller energy corrections. This is illustrated by the natural orbitals (Fig. 4.2c); they give even smaller energy corrections when allowing single broken pairs.

Hund's rule

As a last observation we note that the largest change in energy occurs when including the seniority-four subspace, and this for all orbital sets in Fig. 4.2.

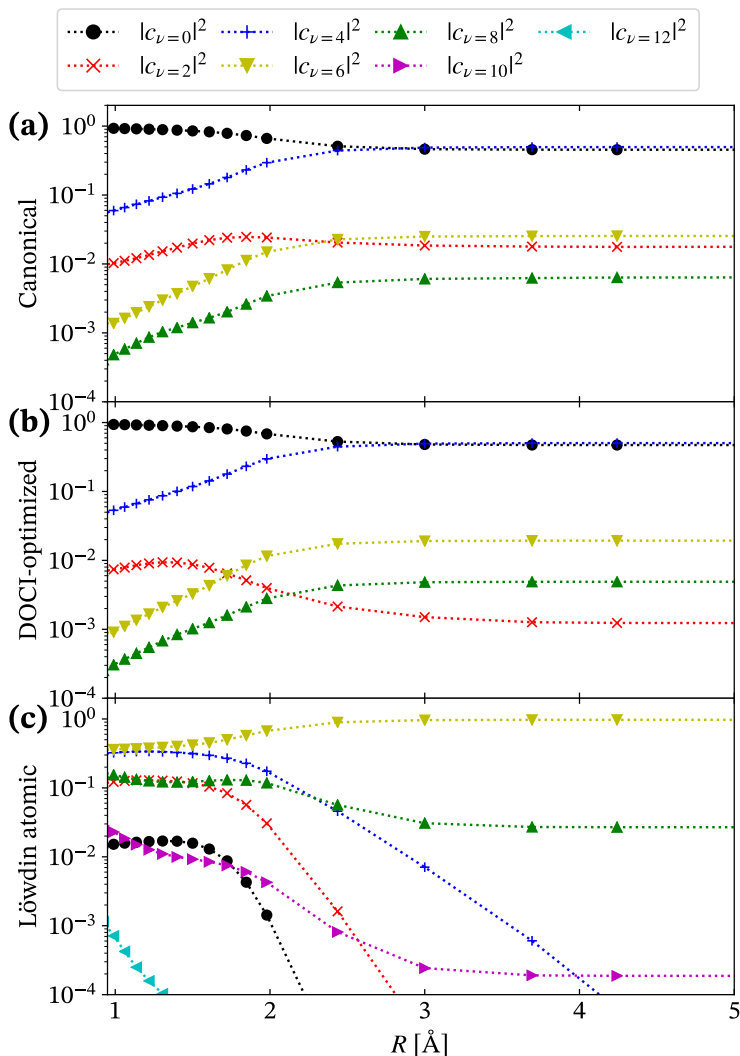


Figure 4.3: Weights of the different seniority subspaces for the ground state wave function of the nitrogen dimer. Results for canonical orbitals (a), DOCI-optimized orbitals (b) and Löwdin orthogonalized atomic orbitals (c) are given.

This trend was also noticed in Ref. [226] for the nitrogen dimer in nonlocal orbitals. When including determinants up to seniority-four the energies are close to FCI around the binding distance; however, the binding energy itself is still overestimated due to missing dynamical correlation at the dissociation.

Intuitively, we would expect a much larger error when excluding the seniority-six subspace; Hund's rule dictates dissociation to two nitrogen atoms

with each three unpaired electrons. However, seniority and pairing is an orbital-dependent concept [244]; we need to keep in mind that Hund's rule applies to a nitrogen atom with orbitals localized around that atom. To study the interpretation of Hund's rule in non-local orbitals, we consider a toy model of two sets of three orbitals (p_x, p_y, p_z) and (p'_x, p'_y, p'_z) . Each set of orbitals mimics the local p -orbitals of each nitrogen atom, which are singly occupied and couple together to a $S = 3/2$ state, as dictated by Hund's rule. Our performed tensor network calculations target a singlet state for the dimer, so the two toy-nitrogen atoms should couple as $[3/2, 3/2]^0$. Mimicking non-local orbitals, we rotate the orbitals pairwise as follows:

$$\begin{aligned} \pi_1 &= p_x \cos \theta + p'_x \sin \theta, & \pi_1^* &= -p_x \sin \theta + p'_x \cos \theta \\ \pi_2 &= p_y \cos \theta + p'_y \sin \theta, & \pi_2^* &= -p_y \sin \theta + p'_y \cos \theta \\ \sigma &= p_z \cos \theta + p'_z \sin \theta, & \sigma^* &= -p_z \sin \theta + p'_z \cos \theta. \end{aligned}$$

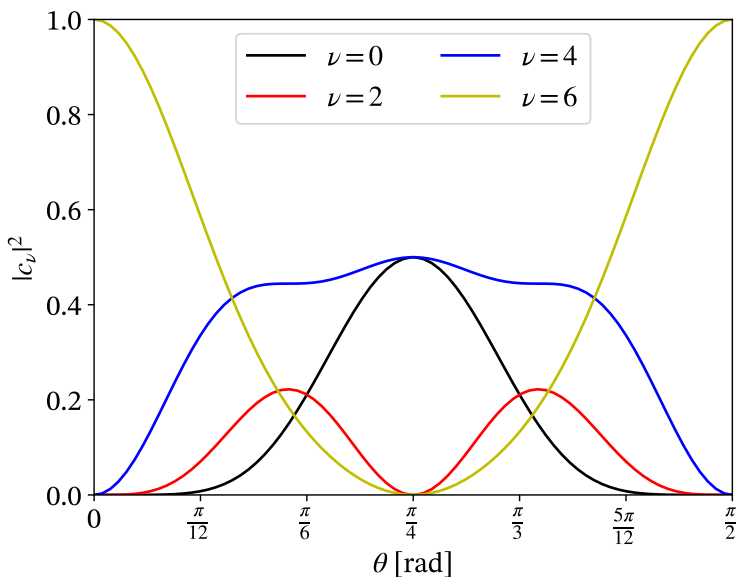


Figure 4.4: Toy model of two nitrogen atoms with both $S = 3/2$ respecting Hund's rule. The two atoms couple together to a singlet. The figure represents the weights of the different seniority sectors for local $(0, \pi/2)$, delocalized $(\pi/4)$ orbitals and everything in between.

In Fig. 4.4, the weights of the different seniority sectors are given for the $[3/2, 3/2]^0$ coupled toy wave function in function of the rotation angle θ . As can be seen in this model seniority-six is actually of no importance when working

with delocalized orbitals ($\theta = \pi/4$). Instead, the correct dissociation can be described with only seniority-zero-plus-four and both seniorities equally important.

Both canonical orbitals and the DOCI-optimized orbitals are delocalized for the $2p$ -orbitals in this system. The dominating importance of the seniority-zero and four for the wave function at dissociation is very clear in Fig. 4.3a and Fig. 4.3b. The other seniority sectors have very small contributions in comparison. As an illustration, we also included calculations with Löwdin orthogonalized atomic orbitals in Fig. 4.3c. As these orbitals are localized, it corresponds with $\theta = 0$ in Fig. 4.4. These orbitals do give rise to a very important seniority-six subspace at dissociation, as predicted by Hund's rule. Furthermore, all seniority sectors smaller than six express a superexponential decay.

4

4.2.2. Benzene

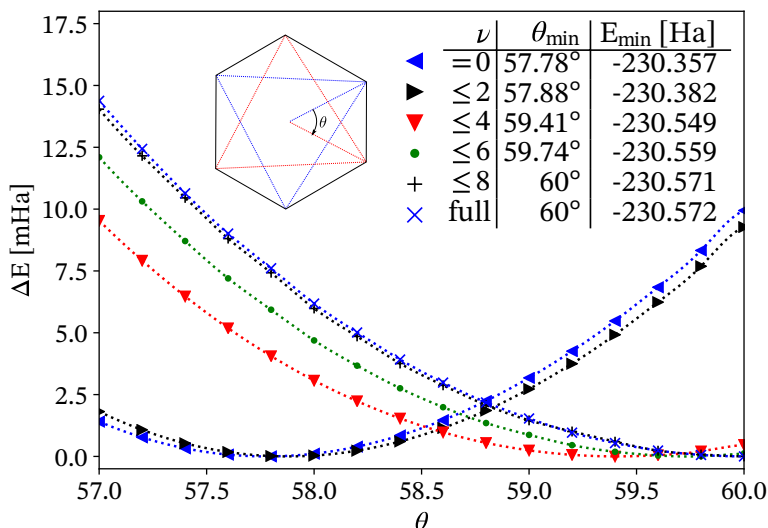


Figure 4.5: Benzene in a STO-6G basis set for different distortion angles. The minimal energies and the corresponding distortion angles for increasing seniority subspaces are given in the inset. A graphical depiction of the in-plane distortion of the aromatic ring in benzene is also shown.

In this section, the in-plane distortion of benzene is studied. The exact nature of the distortion is given in the inset in Fig. 4.5 and is characterized by the angle θ . At $\theta = 60^\circ$ the D_{6h} symmetric ground state geometry of benzene is obtained. At other angles, the distortion introduces alternating shorter and longer carbon-carbon bonds. For this system Boguslawski *et al.* [233] showed that benzene ($\theta = 60^\circ$) is not the ground state configuration within the seniority-zero subspace; an artificial symmetry breaking occurs when al-

lowing orbital optimization.

We use DOCI-optimized orbitals in the STO-6G basis set to study this artificial symmetry breaking with all electrons correlated. We chose STO-6G as the distortion angle of the minimal energy DOCI structure is particularly large for this basis set. The tensor network calculations are executed with a reduced bond dimension of 1000.

In Fig. 4.5 the results for the ground state in the different seniority subspaces are given. In accordance with Boguslawski et al. [233], we notice that, indeed, the ground state is not found at $\theta = 60^\circ$ in the seniority-zero subspace. When the breaking of one pair is allowed in this orbital set, the correction is rather small and the correct symmetry is not restored. This is expected due to the aforementioned first order decoupling of the seniority-zero and seniority-two subspace in these orbitals.

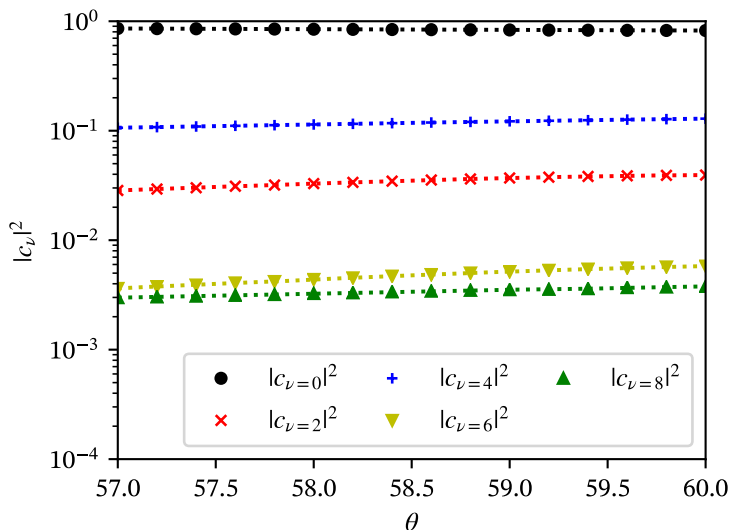


Figure 4.6: Weights of the different seniority subspaces for the ground state wave function of benzene under in-plane distortion. Results are given for DOCI-optimized orbitals.

When including progressively higher seniorities, the stable configuration moves closer to the expected D_{6h} -symmetric benzene. The potential energy surface enjoys a large qualitative correction when including the seniority-four subspace in the calculations; however the minimum of the potential energy surface is still off by 0.59° . The inclusion of seniority-six further improves the quality of the potential energy surface, but only at seniority-eight the exact ground state seems to be recovered, at least up to the resolution of our per-

formed calculations. At this point, the results become very close to the untruncated seniority results. In Fig. 4.6, the weights of the different seniority sectors in the ground state during distortion are also given. These weights do not express the large changes as seen during dissociation of the nitrogen dimer in Fig. 4.3. This is quite expected; the bond breaking of the nitrogen dimer is a far more outspoken change than the small benzene distortions in this section.

4.2.3. Neon dimer

The neon dimer, constituting out of two noble gas atoms, is very weakly bonding. Although the electrons do not feel the need to form covalent bonds between the two atoms, it expresses some bonding character due to weak dispersion forces. In Ref. [245], an empirically fitted potential curve results in a binding energy of $-134 \mu E_h$ and a binding distance of 3.091 \AA .

As the binding of the neon dimer is rather weak and is due to dynamical correlations, it will be very sensitive to the chosen basis set size. For an accurate description of the potential energy curve, a large basis set should be chosen and basis set superposition errors (BSSE) should be taken into account appropriately [246]. A clear example of the importance of BSSE-corrections is the dissociation curve on the Hartree-Fock level. At this level of theory no binding is expected as the Hartree-Fock solution is dispersion-free. However, when using small basis sets, one would find a binding neon dimer at the Hartree-Fock level if one neglects to correct the BSSE [246].

We study the neon dimer in the aug-cc-pVDZ basis, which was found to have a favorable tradeoff between mitigating BSSE and numerical stability issues of larger basis sets. Calculations with different allowed seniority sectors are executed while using DOCI-optimized orbitals with a frozen $1s$ core. Reduced bond dimensions up to 800 are used for the DMRG calculations. As the aug-cc-pVDZ basis is a rather small basis for capturing dispersion forces, appropriately removing BSSE is important. This is done by using the Boys and Bernardi counterpoise correction [247].

In Fig. 4.7 the raw uncorrected results are given for the different calculations. For all seniority calculations the neon dimer seems to be bonding. However, for seniority-zero and seniority-two-plus-zero the bonding seems to be very weak; only for $\nu \leq 4$ calculations and higher the bonding character is qualitatively corresponding with the untruncated seniority case.

For the counterpoise correction, equivalent calculations as for the dimer are executed but with one neon atom replaced by a chargeless, electronless ghost atom. This way, we can approximately correct for the extra stabilization each neon monomer feels by the basis functions of the other monomer. The

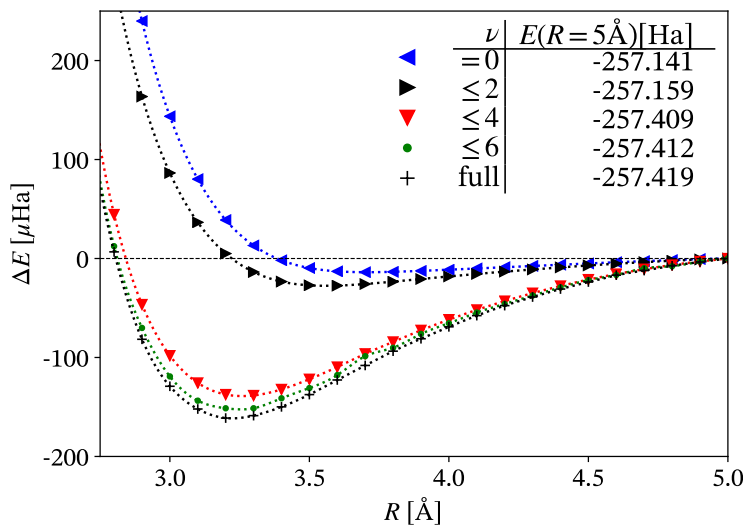


Figure 4.7: Dissociation curves for the neon dimer without BSSE-correction. The energies at large separation distances is given in the inset.

BSSE-corrected dissociation energy for the dimer at distance r is then given by

$$E_{\text{dissoc}}(r) = E_{\text{Ne-Ne}}(r) - E_{\text{Ne-ghost}}(r) - E_{\text{ghost-Ne}}(r). \quad (4.14)$$

The same level of theory should be used for these ghost calculations as for the original calculation. This poses a difficulty since the seniority restricted calculations are not size consistent; $E_{\text{dissoc}}(r \rightarrow \infty)$ does not tend to zero as is desired. Assume we have executed a dimer calculation with $\nu \leq 4$, using ghost calculations with the same $\nu \leq 4$ will overcorrect, while ghost calculations at the lower $\nu \leq 2$ will undercorrect. We try to solve this problem by both over and undercorrecting the calculations and shift both curves to 0 in the dissociation limit. Results for these BSSE-corrections are given for different seniority sectors in Fig. 4.8 with the grayed area indicating where the exact BSSE-correction is expected to be. From Fig. 4.8a and 4.8b, it seems that the weak bonding character present in Fig. 4.7 for $\nu = 0$ and $\nu \leq 2$ practically or completely disappears when taking BSSE-corrections into account. When correcting $\nu \leq 4$ calculations, we can again overcorrect or undercorrect. The overcorrected dissociation underestimates the dissociation energy a bit with respect to the FCI BSSE-corrected calculations while the undercorrected dissociation overestimates the dissociation energy, as can be expected. It seems that calculations with seniority-zero and seniority-two do not model the needed dispersion and at least seniority-four is needed.

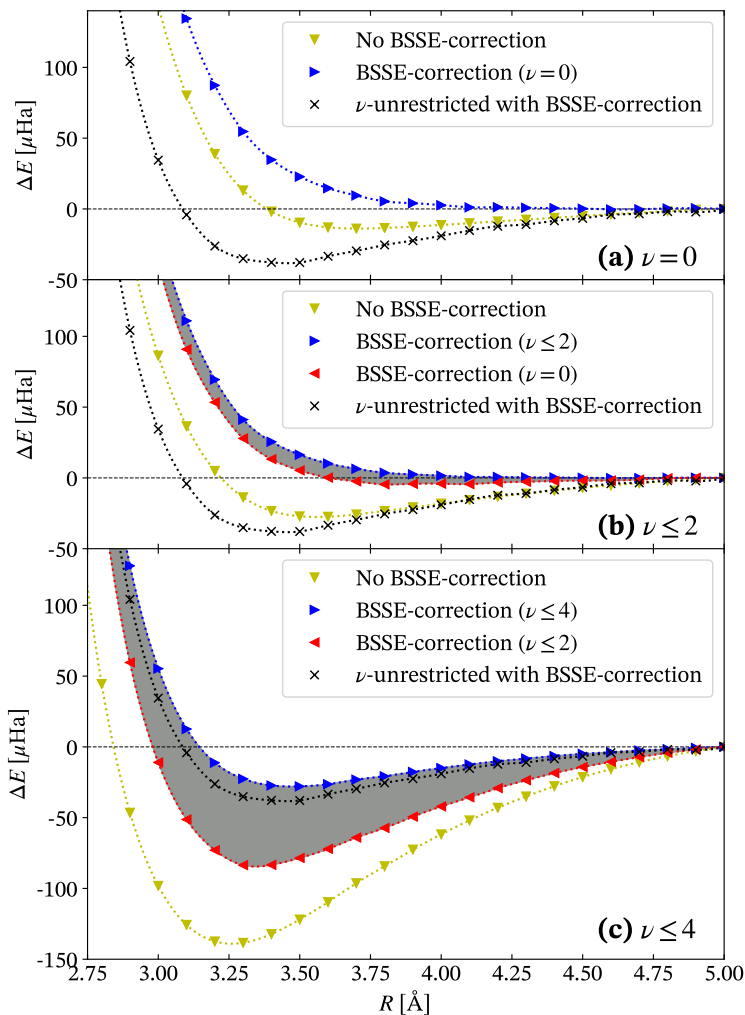


Figure 4.8: Dissociation curves for the neon dimer with and without BSSE correction for different seniority subspaces. When BSSE-correction is performed, the used seniority subspace for the ghost atom calculation is given in brackets in the legend. The dissociation curve for a full-seniority calculation with full BSSE-correction is also shown in each subfigure. Results are shown for $\nu = 0$ (a), $\nu \leq 2$ (b) and $\nu \leq 4$ (c) subspace calculations.

Finally, we notice that the FCI dissociation energy with BSSE-correction is a factor of three smaller than empirical measurements and the bond length is overestimated. This is quite normal when studying dynamical correlations in small basis sets. The limited basis set does not allow all the needed flexibility for the stabilization of the dimer.

4.2.4. A rubidium phenanthrene coordination complex

bond dimension	truncation error	energy [E_h]
canonical orbitals		
25	3.4311×10^{-6}	-6080.794 467
50	7.9219×10^{-7}	-6080.794 988
100	1.5275×10^{-7}	-6080.795 096
150	5.0740×10^{-8}	-6080.795 110
split-localized orbitals		
25	8.2824×10^{-5}	-6081.980 409
50	3.3030×10^{-5}	-6081.984 022
100	3.2201×10^{-6}	-6081.984 583
150	1.3259×10^{-6}	-6081.984 695
200	8.4201×10^{-7}	-6081.984 741
250	4.2780×10^{-7}	-6081.984 759

Table 4.2: DOCI energies for different bond dimensions for the rubidium phenanthrene coordination complex. Calculations in both canonical orbitals and split-localized orbitals were performed. The maximal norm of the discarded singular values during the T3NS optimization is used for the truncation error.

We will conclude this chapter by studying the DOCI-TNS method for large systems. A coordination complex of a Ru^{2+} ion surrounded by three phenanthrene molecules is studied in a minimal basis set (STO-3G). By correlating all electrons this problem results into the optimization of 162 electron pairs in 261 orbitals. This is a particularly hard problem as the system is close to half-filling and it is far larger than possible by QC-DMRG methods working in the full Hilbert space.

In Table 4.2, DOCI-T3NS results are given for both canonical orbitals and split-localized orbitals. Canonical orbitals are often not optimal for DOCI calculations as seen in Section 4.2.1. Only a small amount of correlation can be recovered in this orbital space; as such, low bond dimensions already suffice for an accurate DOCI wave function. This is reflected in the low truncation error even at low bond dimensions. Due to the low bond dimensions needed and the lower computational complexity of DOCI-TNS, the calculations for the canonical orbitals were easily executed on a common laptop.

On the other hand, split-localized orbitals are a good alternative when not explicitly optimizing the orbitals for DOCI; they are localized – a property that is frequently observed in DOCI-optimized orbitals – and since we localized occupied and virtual space of the Hartree-Fock wave function separately, we kept the Hartree-Fock Slater determinant in the DOCI space. Finding the DOCI

wave function in this orbital set poses more of a challenge as much more correlations can be found; we see an improvement of more than $1 E_h$ when comparing with the canonical orbitals. At $D = 250$, 21 sweeps through the tensor network were needed for a convergence of $0.1 \mu E_h$ between subsequent sweeps; each sweep took about 45 minutes on 9 cores. At $D = 150$, each sweep took 5 minutes on 9 cores.

For both orbital sets, we used a radially expanding T3NS network. More information on the orbital ordering can be found in Section 5.3.3.

4.3. A short recap

We joined the concept of seniority with tensor network methods. By using seniority-invariant tensors, we can force all the renormalized states in the virtual bonds to have a well defined seniority number. This allows for arbitrary seniority-restricted calculations. However, there is a possible need for larger bond dimensions than performing regular QC-T3NS calculations. For DOCI (doubly occupied configuration interaction) calculations, we can immediately implement the DOCI-projected quantum chemical Hamiltonian in Eq. (4.11). This results in a fast tensor network calculation, partly because of the simpler Hamiltonian, partly because the entanglement in the seniority-zero subspace is less substantial than for seniority-unrestricted calculations. Large scale DOCI-TNS calculations were shown in Section 4.2.4. The DOCI-TNS algorithm provides a fast initialization for subsequent QC-TNS calculations.

Several systems are studied within different seniority subspaces. For the dissociation of the nitrogen dimer, only a quantitative dissociation curve can be obtained when at least two pairs are allowed to be broken. This can be theoretically explained due to Hund's rule. The planar distortion of benzene and its artificial D_{6h} symmetry breaking in the seniority-zero subspace [233] is also studied. A large correction of the artificial symmetry breaking occurs when including seniority-four, however up to eight unpaired electrons are needed for a complete restoration of the correct benzene ground state in the used basis set. We also considered the dissociation of the neon dimer. At the seniority-zero level of theory the neon dimer is non-binding; DOCI does not capture the dispersion forces needed for the weak binding characteristic of neon. Only at seniority-four and onward, the dispersion forces are adequately picked up.

For all systems, the seniority-two subspace has only a small contribution to the total wave function when using DOCI-optimized orbitals; this is expected by the theoretical first order decoupling between seniority-zero and seniority-two subspaces in these types of orbitals [232]. However, a first order decoupling is not an exact decoupling and other orbital sets can be found which attribute even less importance to the seniority-two subspace. An example of

this is given by the natural orbitals of the nitrogen dimer in Fig. 4.2c.



5

Orbitals in tensor networks

It is much more rewarding to do more with less.

Donald Knuth

In the previous chapters we extensively discussed the T3NS algorithm. Among others, we discussed the usage of symmetries of the chemical Hamiltonian in the algorithm. Although this provides a faster and more accurate optimization, it cannot perform miracles. If the entanglement structure of the ground state is badly represented by the tensor network, a large bond dimension will be needed for accurate results. A bad representation of the entanglement structure can originate from a bad orbital choice (discussed in Section 5.1), a bad orbital ordering on the tensor network (discussed in Section 5.2 and Section 5.3), or a more fundamental problem with the geometry of the network itself. Only the first two origins of a bad entanglement structure can be addressed in the MPS, while the T3NS exhibits some flexibility in its geometry that can be optimized (discussed in Section 5.4).

Compared to DMRG, the T3NS has a less favorable scaling with respect to the chosen bond dimension. As such, it is even more important for the T3NS to get the orbital choice and ordering right. Compensation of a bad orbital choice and ordering by increasing the bond dimension is less of an option for the T3NS.

5.1. The orbital choice

Tensor network methods are not invariant under orbital rotations, although they become less sensitive to the orbital choice with increasing bond dimension. A good choice of orbitals can have a sizable effect on the accuracy and convergence speed of tensor network calculations. For large or elongated molecules the usage of localized orbitals has been shown advantageous for highly accurate DMRG calculations [94, 145, 150, 182, 185, 187, 192, 203, 204, 248–261]. This is not completely surprising as DMRG stems from the notion of locality; it is particularly successful in studying systems with entanglement arising from local interactions between the physical degrees of freedom.

The recent development of the sliced basis density matrix renormalization group (SBDMRG) [262, 263] tries to further exploit locality by slicing the three-dimensional space along the z direction. While the basis states in the x and y directions are still derived from Gaussian basis sets, the z direction becomes discretized. The SBDMRG has proven very useful for linear molecules such as long hydrogen chains.

Exploiting locality along all three dimensions for quantum chemical systems has recently been demonstrated by using multisliced gausslets [264] in junction with DMRG (MSG-DMRG) [265]. Very accurate energies for H_{10} in the complete basis set limit were obtained at low bond dimensions. Gausslets are given by a sum of simple Gaussians centered around different points and are smooth and localized while still keeping the Coulomb interactions diagonal (i.e. the Coulomb interactions are two-point interactions).

Another recent development for finding orbital sets where targeted states have low entanglement has been made by Krumnow *et al.* [189, 266, 267]. Rather than constructing a spatially local basis set, they opted for optimizing the spatial orbitals within a chosen basis set through Gaussian mode transformations. During the two-site DMRG optimization, the orbitals of the two-site tensor are mixed by rotation to minimize the entanglement in the two-site tensor.

In this chapter we do not focus on choosing a right orbital set, but rather on ordering the chosen orbital set in a satisfactory fashion onto the network.

5.2. Ordering the orbitals

The choice of orbitals and their ordering is of key importance for an efficient and accurate tensor network method in quantum chemistry. In condensed matter physics, one can use the geometry of the lattice as a clear guiding tool for the ordering of the physical degrees of freedom. This obvious choice is lacking in quantum chemistry.

Several methods have been developed for finding good orderings of the orbitals onto the MPS, most of which are easily adapted to the T3NS. Most methods are based on minimizing the bandwidth of certain two-orbital measures A_{ij} , where high values of A_{ij} indicate highly interacting orbitals. A cost function of the form

$$\hat{A}_\eta = \sum_{ij} A_{ij} d_{ij}^\eta \quad (5.1)$$

can be defined, where i and j are the indices of two orbitals (sites) in the tensor network and d_{ij} is the network distance between the orbitals. It is given by the number of edges traversed when going from orbital i to orbital j and is well defined due to the absence of loops in both the T3NS and the MPS. The cost function \hat{A}_η should be minimized for positive η -values, while it should be maximized for negative η -values.

Choosing η equal to two is perhaps the most advantageous, as it allows an approximate solution through the Fiedler vector [149, 268]. We can define a Laplacian $L_A = D - A$ of the matrix A where D is a diagonal matrix with elements $D_{ii} = \sum_j A_{ij}$. The Fiedler vector is then given by the second eigenvector of L_A and is a solution of the constrained minimization problem for

$$A(x) = x^\dagger L_A x = \sum_{ij} A_{ij} |x_i - x_j|^2 \quad (5.2)$$

with $\sum_i x_i = 0$ and $\sum_i x_i^2 = 1$ as constraints. By ordering the x_i values of the Fiedler vector in descending or ascending order, a mapping of the orbitals can be proposed onto a linear MPS. It provides an approximate solution for the minimum of Eq. (5.1) with $\eta = 2$.

For the T3NS the approximate solution provided by the Fiedler vector is not an option, instead we opted to minimize cost functions given in Eq. (5.1) by simulated annealing. We iteratively permute two randomly picked orbitals on the network and accept the given permutation with a probability

$$p = \min(e^{-\beta \Delta \hat{A}_\eta}, 1). \quad (5.3)$$

Here, $\Delta \hat{A}_\eta$ is the difference in the cost function after and before the permutation and β is a parameter tuning the acceptance of permutations that worsen the cost function. This technique was also used by Rissler *et al.* [186] for the optimization of a mutual information-based cost function (see further).

Several two-orbital measures for the cost function in Eq. (5.1) can be proposed. The most straightforward measures to use are the ones based on the one-electron and two-electron integrals in the Hamiltonian. In Ref. [204] the

bandwidth of the T_{ij} one-electron integrals was minimized, effectively grouping orbitals belonging to the same irreps. Moritz *et al.* [248] compared the performance of minimizing the bandwidth for the Coulomb integrals $J_{ij} = V_{ijij}$, the exchange integrals $K_{ij} = V_{ijji}$, and the mean-field matrix $M_{ij} = 2J_{ij} - K_{ij}$. By testing different orderings on the chromium dimer, they concluded that ordering based on the exchange matrix worked best. Minimization of the bandwidth in Refs. [204, 248] was not done by minimizing Eq. (5.1), instead a reverse Cuthill-McKee (RCM) algorithm [269, 270] was used. When using RCM, a cutoff should be defined for the matrix elements of the considered measure as it only permutes the rows and columns of a sparse matrix to move all nonzero elements as close as possible to the diagonal. The Fiedler vector provides a better minimization of the bandwidth as it treats the different magnitudes of the interactions appropriately.

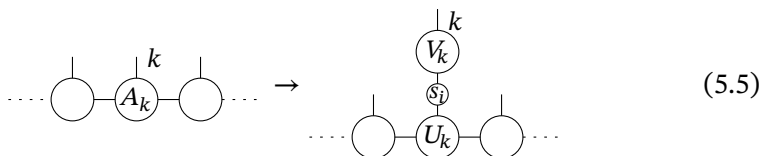
Currently, the exchange matrix is the most widely used Hamiltonian-based measure for localized orbitals [169, 176, 254, 258, 260, 263] as it also reflects the overlap between the orbitals. An ordering by using the Fiedler vector or a genetic algorithm in junction with the exchange matrix is implemented in open-source QC-DMRG projects such as CheMPS2 [92, 144, 176, 193] and Block [148, 187, 199, 204, 258]. Other measures that have been used or proposed are, for example, the Fock matrix for non-HF orbitals [271] and

5

$$\frac{V_{iijj}^2}{\epsilon_i - \epsilon_j}, \quad (5.4)$$

a measure based on the Møller-Plesset perturbation theory [272].

Not only measures based on the Hamiltonian can be used; entanglement-based measures are perhaps even more suitable as the workings of tensor network methods are rooted in entanglement. Legeza and Sólyom [185] proposed the usage of the single-orbital von Neumann entropy. It is the entropy between one orbital and the rest of the system. For T3NS and DMRG, these entropies are quite easily calculated. One only has to bring the network in the canonical form with the appropriate site chosen as orthogonality center. A singular value decomposition along the physical edge immediately results in the needed singular values, i.e.



$$\dots \circlearrowleft A_k \circlearrowright \dots \rightarrow \dots \circlearrowleft \begin{array}{c} V_k \\ S_j \\ U_k \end{array} \circlearrowright \dots \quad (5.5)$$

The single-orbital entropy $S(k)$ is now easily calculated by using Eq. (1.8):

$$S(k) = - \sum_i s_i^2 \ln s_i^2. \quad (5.6)$$

The single-orbital entropy is a measure of the entanglement between the orbital and the rest of the system; as such, Legeza and Sólyom [185] proposed to group the orbitals with the highest entropies together. As there is no *a priori* knowledge of the single-orbital entropies, an initial cheap low-D DMRG calculation was proposed to obtain a first estimate of $S(k)$.

A more suitable entanglement measure is given by the mutual information I_{kl} . This measure, proposed by Rissler, Noack and White [186] for the ordering of orbitals, quantifies the correlations between the two orbitals k and l and is given by

$$I_{kl} = (S(k) + S(l) - S(k, l))(1 - \delta_{k,l}), \quad (5.7)$$

where $S(k, l)$ is the von Neumann entropy between orbitals k and l , and the rest of the system. Although not as straightforwardly obtained as single-orbital entropies, these two-orbital entropies can still be calculated quite efficiently in an MPS and T3NS [186].

Just as for the single-orbital entropies, the calculation of the mutual information needs an initial cheap calculation for the knowledge of the entropies. Luckily, Rissler *et al.* [186] noticed no substantial changes for the mutual information between DMRG calculations at low bond dimension and more accurate calculations. They also noticed that the mutual information between orbitals of the same irrep were particularly large. This further strengthened the already existing heuristic of grouping orbitals per irrep on the MPS, as T_{ij} terms vanish for $I_i \neq I_j$ [149, 204, 273]. Since the introduction of the mutual information, this measure has been widely used for the ordering of orbitals in DMRG [100, 146, 149, 168, 170, 192, 274]. The mutual information has also been used in DMRG for the efficient reconstruction of the CI expansion [275].

The mutual information is generally formulated by using the von Neumann entropy; as such it is a positive value for all k and l due to the subadditivity of the von Neumann entropy [276]. If one would derive mutual informations with respect to the Rényi entropies, the positivity of the mutual information would only be ensured for $\alpha > 1$ as the Rényi entropy is only subadditive for these values [277, 278].

5.3. Reordering the orbitals

Instead of optimizing the orbitals by Gaussian mode transformations as done by Krumnow *et al.* [189, 266, 267], we will opt to only optimize the ordering of

a chosen orbital set onto the tensor network. Optimizing the global ordering of the orbitals can be easily done by repeated local permutations while sweeping through the network. Let us first clarify these local permutations.

5.3.1. The local permutation

Suppose we have a two-site tensor built from two contracted physical sites; we can easily swap the order of the two physical degrees of freedom within the two-site tensor by a permutation of the indices of the tensor, i.e.

$$\frac{a}{\alpha} \textcircled{A} \textcircled{B} \frac{b}{\beta} = \frac{a}{\alpha} \textcircled{AB} \frac{b}{\beta} = \frac{b}{\alpha} \textcircled{AB'} \frac{a}{\beta}, \quad (5.8)$$

where a and b are the physical edges, α and β are the unconnected virtual edges and 1 is a virtual internal edge.

For the T3NS, swapping of the orbitals can be done in a similar fashion. However, swapping of orbitals connected through an intermediate branching tensor is also needed to ensure a global reordering; we need three-site tensors. For example, we can swap the physical degrees of freedom a and b in the following example:

$$\begin{array}{c} \beta \\ | \\ \textcircled{B} \\ | \\ 2 \\ | \\ \textcircled{C} \\ | \\ 1 \\ | \\ \textcircled{A} \\ | \\ \alpha \end{array} \frac{b}{\beta} \frac{a}{\alpha} \gamma = \begin{array}{c} \beta \\ | \\ \textcircled{ABC} \\ | \\ \alpha \end{array} \frac{b}{\beta} \frac{a}{\alpha} \gamma = \begin{array}{c} \beta \\ | \\ \textcircled{ABC'} \\ | \\ \alpha \end{array} \frac{a}{\beta} \frac{b}{\alpha} \gamma. \quad (5.9)$$

Again, the loose virtual edges are denoted by the Greek letters α , β and γ and the virtual internal edges are given by the numbers 1 and 2. Similarly, one could permute all three physical degrees of freedom neighboring a branching site at once by forming a four-site tensor.

For fermionic tensors, the fermionic sign has to be taken into account when permuting the indices of the rank-four tensor in Eq. (5.8) or rank-five tensor in Eq. (5.9). A change of the coupling scheme is also needed if the tensors are additionally written as SU(2)-invariant tensors. As these extra complications were already extensively discussed in Chapter 3, we will only briefly mention the prefactors needed for the two permutations given in Eq. (5.8) and Eq. (5.9). For the first permutation given in Eq. (5.8) we need to deform

$$\frac{a}{\alpha} \textcircled{AB} \frac{b}{\beta} = \frac{x_a}{x_\alpha} \textcircled{AB'_{x_1}} \frac{x_b}{x_\beta} \frac{x_\alpha}{x_\alpha} \frac{x_\beta}{x_\beta} \frac{x_a}{x_\alpha} \frac{x_b}{x_\beta} \quad (5.10)$$

to

$$\frac{b}{\alpha} \overline{AB'} \beta = \frac{x_b}{x_\alpha \alpha_\alpha} \overline{AB''}_{x_2} \frac{x_a}{x_\beta \alpha_\beta} \begin{matrix} \mathbf{x}_b \\ \downarrow \\ \mathbf{x}_\alpha \\ \mathbf{x}_2 \end{matrix} \begin{matrix} \mathbf{x}_a \\ \downarrow \\ \mathbf{x}_\alpha \\ \mathbf{x}_\beta \end{matrix} \quad (5.11)$$

resulting into

$$\frac{x_b}{x_\alpha \alpha_\alpha} \overline{AB''}_{x_2} \frac{x_a}{x_\beta \alpha_\beta} = [j_2]^2 \sum_{j_1} (-1)^{j_1+j_2-j_a-j_b} \begin{Bmatrix} j_a & j_\beta & j_2 \\ j_b & j_\alpha & j_1 \end{Bmatrix} \quad (5.12)$$

$$(-1)^{\pi_a \pi_b} \frac{x_a}{x_\alpha \alpha_\alpha} \overline{AB''}_{x_1} \frac{x_b}{x_\beta \alpha_\beta} .$$

The needed prefactor is obtained by using the 6j-swap property of the symmetry tensor and including the appropriate fermionic sign. For the permutation given in Eq. (5.9) we need to deform

5

$$\begin{matrix} \beta & b \\ \overline{ABC} \\ a & \alpha & \gamma \end{matrix} = \begin{matrix} x_\beta \alpha_\beta & x_b \\ \overline{ABC''}_{x_1 x_2} \\ x_a & x_\gamma \alpha_\gamma \\ x_\alpha \alpha_\alpha \end{matrix} \begin{matrix} \mathbf{x}_\beta \\ \mathbf{x}_b \\ \mathbf{x}_a \\ \mathbf{x}_\alpha \\ \mathbf{x}_\gamma \\ \mathbf{x}_2 \end{matrix} \quad (5.13)$$

to

$$\begin{matrix} \beta & a \\ \overline{ABC'} \\ b & \alpha & \gamma \end{matrix} = \begin{matrix} x_\beta \alpha_\beta & x_a \\ \overline{ABC''}_{x_3 x_4} \\ x_b & x_\gamma \alpha_\gamma \\ x_\alpha \alpha_\alpha \end{matrix} \begin{matrix} \mathbf{x}_\beta \\ \mathbf{x}_a \\ \mathbf{x}_b \\ \mathbf{x}_\alpha \\ \mathbf{x}_\gamma \\ \mathbf{x}_4 \end{matrix} \quad (5.14)$$

resulting into

$$\begin{matrix} x_\beta \alpha_\beta & x_a \\ \overline{ABC''}_{x_3 x_4} \\ x_b & x_\gamma \alpha_\gamma \\ x_\alpha \alpha_\alpha \end{matrix} = [j_3]^2 [j_4]^2 \sum_{j_1 j_2} (-1)^{j_2+j_b-j_4-j_a} \begin{Bmatrix} j_\alpha & j_a & j_1 \\ j_b & j_\beta & j_2 \\ j_3 & j_4 & j_\gamma \end{Bmatrix} \quad (5.15)$$

$$(-1)^{\pi_a \pi_2 + \pi_\beta \pi_b} \begin{matrix} x_\beta \alpha_\beta & x_b \\ \overline{ABC''}_{x_1 x_2} \\ x_a & x_\gamma \alpha_\gamma \\ x_\alpha \alpha_\alpha \end{matrix} ,$$

which is obtained by using the 9j-swap and the reflection property of the symmetry tensor and again including the appropriate fermionic sign. Prefactors for other permutations of orbitals around a branching tensor can be found analogously.

5.3.2. Optimizing the orbital ordering

There are two options for optimizing the global order through local permutation of orbitals. In the first, one can permute during the energetic optimization of the wave function; every time an optimized multi-site tensor has been obtained, the best permutation of its local degrees of freedom can be searched. Intertwining orbital reordering with energy optimization has the advantage that one can directly permute the *untruncated* multi-site tensor; however, two-site optimization is insufficient when treating a T3NS in this case and one has to opt for three-site or four-site optimization to ensure a global permutation of the local physical degrees of freedom. Another disadvantage is that the renormalized operators have to be appropriately adapted at each permutation.

In the second, the optimization of the orbital ordering is done separately from the energetic optimization; a sweeping procedure designated for orbital reordering is executed once an optimized wave function is obtained. This procedure has the advantage that the renormalized operators can be completely disregarded during the reordering; the operators can be recalculated once a new order is obtained. This procedure has been implemented in our T3NS code [156] and is called by using the `--disentangle` flag.

Selecting the best order

The most straightforward way of choosing the best permutation would be to accept the ordering for which the truncated SVD of the multi-site tensor gives the best approximation. However, this is only an option when orbital ordering and energy optimization are intertwined and not when optimizing the order in a separate sweep. For the latter case, the multi-site tensor to decompose consists out of a contraction of several single-site tensors with a given bond dimension D . As such, any truncated decomposition at the same bond dimension will be error-less for the original order.

Alternatively, entropy measures such as the von Neumann entropy and the Rényi entropy can be used for determining the best order. These entropy-measures were introduced in Section 1.1.1 and are respectively given by

$$S = - \sum_i s_i^2 \ln s_i^2 \quad (5.16)$$

and

$$S_\alpha = \frac{1}{1-\alpha} \ln \left(\sum_{i=0}^m s_i^{2\alpha} \right) \quad (5.17)$$

where s_i are the singular values of an SVD. We remind the reader that the von Neumann entropy corresponds to the Rényi entropy for $\alpha \rightarrow 1$.

When reordering the orbitals on two neighboring sites as in Eq. (5.8), the entropy of the two different singular value decompositions can be compared. When reordering the physical edges as in Eq. (5.9), multiple SVDs are needed for the decomposition and each SVD has an associated entropy. We opted for comparing the sum of these entropies in this case; other options are also possible such as comparing the maximal value of the entropies.

As the permutations are local, it is quite sensitive to local minima during optimization. We partially remedy this by accepting suboptimal local permutations with exponentially decaying chance; i.e. we accept a certain permutation with a probability $p = \min(e^{-\beta \Delta S_\alpha}, 1)$ where ΔS_α is the difference in entropies after and before permutation. The parameter β can be freely chosen and tunes the willingness to accept permutations with a higher entropy.

5.3.3. Numerical examples

In this section, some numerical results for the reordering of the orbitals onto an MPS and a T3NS are presented. We will compare ordering minimizing the cost function \hat{K}_2 , i.e. the cost function in Eq. (5.1) for the exchange matrix at $\eta = 2$, with results after local permutations. Both the von Neumann entropy (S_1) and the Rényi entropy with $\alpha = 1/4$ ($S_{1/4}$) are used for the latter. The ordering with \hat{K}_2 as cost function is optimized by the Monte Carlo-esque acceptance given in Eq. (5.3) with $\beta = 1$.

Local permutations are accepted with a probability $p = \min(e^{-\beta \Delta S_\alpha}, 1)$ where $\beta = 20$ for both S_1 and $S_{1/4}$. After some experimentation, this value provided reliable convergence to lower entangled orderings while still allowing suboptimal permutations to move out of local minima. All calculations were converged up to an energy difference of 10 nE_h between subsequent sweeps.

The $[\text{Cu}_2\text{O}_2]^{2-}$ bisoxo isomer

As a first case, we will revisit the $[\text{Cu}_2\text{O}_2]^{2-}$ bisoxo isomer. We use the Stuttgart pseudopotential and associated basis functions (ECP10MDF) [201] for Cu and the atomic natural orbital (ANO) basis set of Pierloot *et al.* for O [202]; the same active space was used as in Refs. [104, 149]. In Table 5.1, some results are summarized for several DMRG calculations on said system. We exploited

both the D_{2h} point group symmetry and $SU(2)$ -symmetry in the following calculations; the reported bond dimensions D_{red} are *reduced* bond dimensions.

	order	D_{red}	Energy [E_h]
(1)	Exchange matrix	200	-541.527 808
(2)	Reordering (1) by von Neumann entropy	200	-541.532 380
(3)	Reordering (1) by Rényi entropy ($\alpha = 1/4$)	200	-541.533 353
(4)	Reordering (1) by Rényi entropy ($\alpha = 1/4$)	1000	-541.538 769
(5)	Exchange matrix	1000	-541.538 469
(6)	Reordering (5) by Rényi entropy ($\alpha = 1/4$)	1000	-541.539 167
(7)	Random ordering	1000	-541.535 873
(8)	Reordering (7) by Rényi entropy ($\alpha = 1/4$)	1000	-541.537 294
(9)	Reordering (8) by Rényi entropy ($\alpha = 1/4$)	1000	-541.537 340
(10)	Reordering (9) by Rényi entropy ($\alpha = 1/4$)	1000	-541.537 370

Table 5.1: Results for DMRG calculations of the $[\text{Cu}_2\text{O}_2]^{2-}$ bisoxo isomer. The extrapolated result for the bisoxo isomer was given in Table 3.2 to be $-541.540\,12\,E_h$.

Optimizations at a reduced bond dimension of 200 and 1000 have been executed where the orbital order was determined by the exchange matrix; these are calculations (1) and (5) in Table 5.1. Calculations (2) and (3) use the converged calculation (1) and reorder the orbitals with the von Neumann entropy and the Rényi entropy with $\alpha = 1/4$ as criterion, respectively. Afterwards the optimization is continued at $D_{\text{red}} = 200$. Calculation (4) is a continuation of calculation (3) at $D_{\text{red}} = 1000$. Finally, calculation (6) is given by a Rényi entropy based reordering of the wave function of calculation (5) after which optimization is continued.

The results enumerated in Table 5.1 make it quite clear that optimizing the order by local permutations improves upon the initial order obtained through the exchange matrix. For both $D_{\text{red}} = 200$ and 1000, the error with the extrapolated result is approximately halved after reordering with the Rényi entropy at $\alpha = 1/4$. Using the von Neumann entropy as criterion for the local permutations also improves upon the initial ordering in (1) but to a lesser extent than the Rényi entropy. Although no thorough research has been done to the optimal α -value for the entropy and the optimal β -value for the acceptance of suboptimal permutations, we noted consistently better performances for $S_{1/4}$ than for S_1 ; this is probably due to the heavier weight assigned to slowly decaying tails of singular values for the former measure.

In Fig. 5.1, a heat map of the Rényi entropy at all α -values and of the singular values is given for calculations (5) and (6) for each virtual bond in the MPS. No large qualitative changes can be noted in the entropy-profile of the

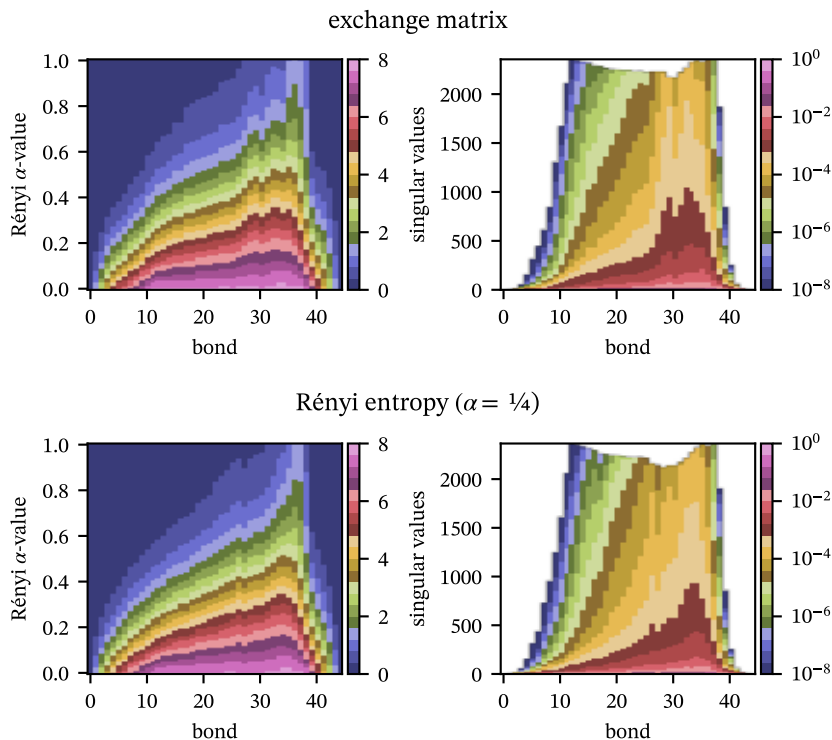


Figure 5.1: Heat map of the Rényi entropy and the singular values along each bond in the MPS with $D_{\text{red}} = 1000$ for the $[\text{Cu}_2\text{O}_2]^{2-}$ bisoxo isomer. The orbitals were ordered by the cost function \hat{K}_2 using the exchange matrix in Eq. (5.1) (upper) and by subsequently optimizing the order using the Rényi entropy at $\alpha = 1/4$ (lower). These wave functions correspond with (5) and (6) in Table 5.1.

MPS before and after optimizing the orbital order; however, the profile after optimizing seems more smoothed (especially around bonds 30 to 40) and the entropies along the bonds are also lowered, resulting in more accurate calculations.

As an extra illustration, we have performed calculations with an initial random ordering (denoted as (7) in Table 5.1). As could be expected, these calculations at $D_{\text{red}} = 1000$ perform much worse than the other calculations (4–6). We have subsequently optimized the ordering by local permutations with the Rényi entropy $S_{1/4}$ followed by a new energy optimization at $D_{\text{red}} = 1000$. This has been repeated three times and the results are denoted (8), (9) and (10) in Table 5.1. Although the optimized ordering does perform substantially better than the initial random ordering, it is still worse than the ordering based on

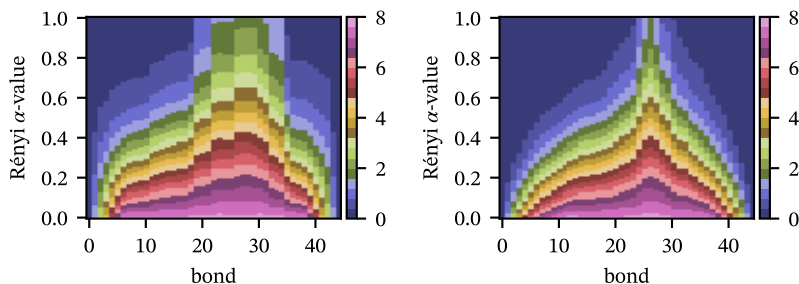


Figure 5.2: Heat map of the Rényi entropy along each bond in the MPS with $D_{\text{red}} = 1000$ for the $[\text{Cu}_2\text{O}_2]^{2-}$ bisoxo isomer. The orbitals were ordered randomly (left) and by subsequently optimizing the order using the Rényi entropy at $\alpha = 1/4$ three times (right). These wave functions correspond with (7) and (10) in Table 5.1.

5

the exchange matrix. The Rényi entropy profiles of the wave function with random ordering (7) and the wave function after reordering trice (10) are given in Fig. 5.2. We note again the same smoothing and lowering of the entropies by local permutations, but again, the overall shape does not change substantially; peaks arise at the same region.

A rubidium phenanthrene coordination complex

As a next exposition of the local reordering, we will study a rubidium phenanthrene coordination complex in a minimal basis set with split-localized orbitals through DOCI-T3NS. This is the same system studied in Section 4.2.4 and has 162 electron pairs in 261 orbitals. A radially expanding T3NS is used with a maximal bond dimension of 150. In Fig. 5.3, heat maps of the Rényi entropies and of the singular values are given along the virtual bonds of the network before and after reordering. While a mapping of the bonds on a linear axis is quite straightforward for an MPS it is not for a T3NS; hence, we have chosen to map the bonds in decreasing order of $S_{1/4}$ entropy.

Although the tensor network is quite big (261 physical sites), most of the virtual legs have a low bond dimension; in fact only 60 of the 390 virtual bonds needed to be truncated. This is visible in the heat maps of the singular values in Fig. 5.3 and is expected as approximately half of the physical sites are at the border of the network for a radially expanding T3NS.

The initial ordering is decided through the exchange matrix. After the optimization of a T3NS with a bond dimension of $D = 150$, the ordering is further optimized by using local permutations with the Rényi entropy $S_{1/4}$ as swapping criterion. The wave function with the newly obtained orbital order is finally further optimized at a bond dimension of 150.

The ground state energies are given by $-6081.983\,891\,E_h$ for the initial or-

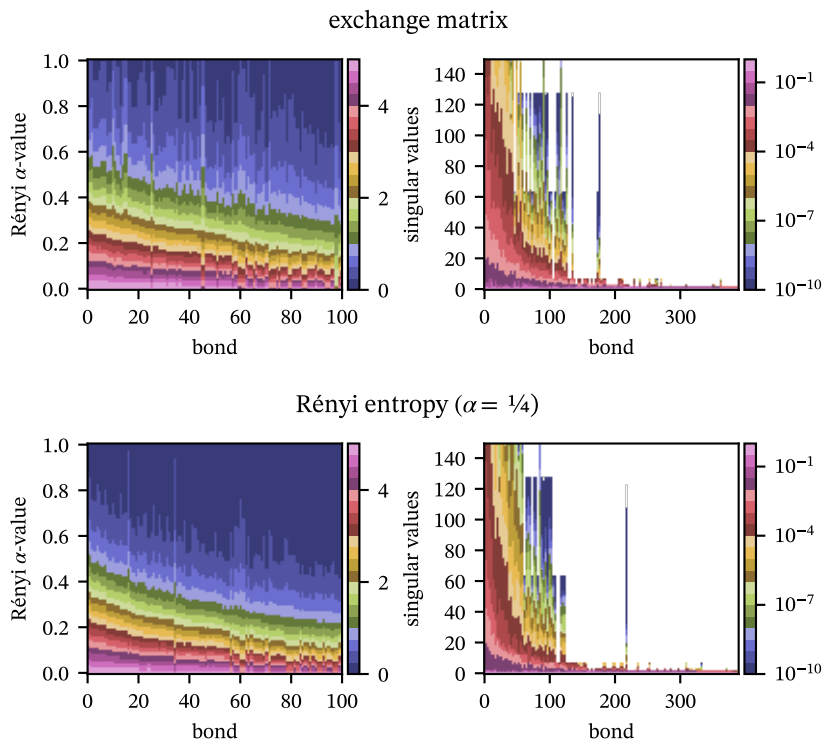


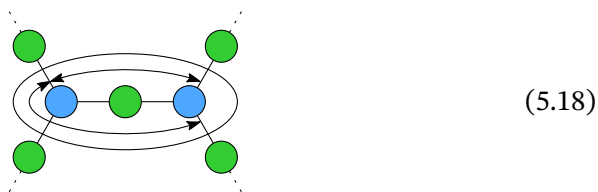
Figure 5.3: Heat map of the Rényi entropy and the singular values along each bond in the T3NS with $D = 150$ for the rubidium phenanthrene coordination complex for a DOCI-T3NS calculation. The bonds are plotted in order of decreasing Rényi entropy at $\alpha = 1/4$. Only the Rényi entropies for the 100 most entangled bonds are shown. The orbitals were ordered by using \hat{K}_2 as cost function (upper) and by subsequently optimizing the order using the Rényi entropy at $\alpha = 1/4$ (lower).

dering and $-6081.984\,695 E_h$ for the final ordering. A notable improvement of the energy is obtained, this is especially apparent when comparing with the reported results in Table 4.2; even calculations at a bond dimension of 50 with the final ordering perform better than the initial exchange matrix-based ordering at $D = 150$.

When looking at Fig. 5.3, it is again quite apparent that reordering of the orbitals through local permutations smooths the entropy profiles and decreases the entropies a bit. We have only plotted the Rényi entropies of the 100 most entangled bonds.

5.4. Deforming the tree tensor network

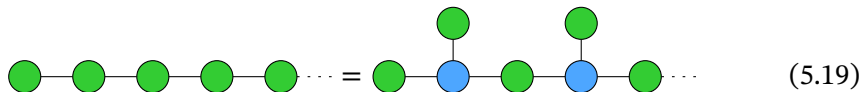
Not only the orbital order can be optimized for a T3NS, but also the geometry of the network itself. Contracting two neighboring tensors and permuting the indices as previously done for the TTNS [279] is not an option for the T3NS as it does not preserve the T3NS structure. Instead, the virtual indices of a multi-site tensor consisting out of two branching and one physical tensor should be permuted; two unique swaps can be identified in this case, i.e.



5

Where physical tensors are colored green while branching tensors are blue. These types of deformations leave the T3NS structure intact; no two branching tensors are adjacent which is elemental for an efficient two-site optimization.

By executing a QR-decomposition on every second physical tensor in an MPS, we can easily deform it as follows



where again the same color scheme is used. Once this form is obtained, the MPS can be changed to a general T3NS by repeated usage of Eq. (5.18).

The implementation of these deformations is a work in progress; hopefully they will shed more light on the competitiveness between T3NS and DMRG as it gives us the opportunity for a controlled transformation between the two.

5.5. A short recap

We have started this chapter with a concise discussion of orbital choices for QC-TNS calculations. Especially the recent research by White with respect to multisliced gausslets [264, 265] seems promising as it exhibits a high degree of locality while still being smooth and keeping Coulomb interactions diagonal.

The largest portion of this chapter deals with the ordering of a chosen orbital set onto the MPS or T3NS. We first discussed existing ordering methods; the most commonly used methods to this day are grouping orbitals with high mutual information or high exchange terms. In Section 5.3 we discussed local permutations of the orbitals on an existing T3NS or MPS. The time needed for locally permuting sites on the tensor network is negligible compared to the

time needed for energetic optimizations while it provides an increased accuracy compared to ordering based on the exchange matrix. As such, it seems to provide an efficient way to further enhance the accuracy of the T3NS.

Local permutations seem to only smoothen the initial ordering. Although it improves the ordering, it does not seem to be capable of finding a more optimal ordering with a vastly different entropy profile than the initial one. Maybe this can be resolved by finding an acceptance of worse orderings which is not too stringent while still improving the final result. Experimenting with, for example, decreasing temperatures $1/\beta$ (as is common in simulated annealing) did not solve the problem.

Bear in mind that these shown results are rather preliminary; research for comparing the ordering by local permutations with other ordering methods (such as the ones based on mutual information) is ongoing. Furthermore, we did not experiment much with the control parameters for the simulated annealing yet.

Finally, the deformations of the tree tensor networks as discussed in Section 5.4 can have a large influence on the accuracy of the T3NS. These types of deformations are not implemented yet, but can allow us for a systematic comparison between the MPS and the T3NS. As the T3NS has an increased flexibility for the representation of entanglement compared to DMRG, it is quite important to use this flexibility as optimal as possible.



6

Conclusions

A story has no beginning or end: arbitrarily one chooses that moment of experience from which to look back or from which to look ahead.

Graham Greene

Since their advent in condensed matter physics, tensor network state methods have known wide success in strongly correlated systems. For example, the density matrix renormalization group (DMRG) provides an efficient algorithm for the study of linear lattice models with local gapped Hamiltonians, while the projected entangled pair states (PEPS) are particularly well-suited for higher-dimensional systems. The geometry of these tensor networks mimics the way in which the different local degrees of freedom are entangled with each other; it allows for an economical parametrization for states with low entanglement.

In quantum chemistry, DMRG has also known substantial success for the simulation of strongly correlated molecules. However, the linear geometry of the matrix product state (MPS), the underlying wave function for DMRG, seems only ideal for encoding the correlations of linear or elongated molecules. As DMRG provides an efficient optimization algorithm, the suboptimal geometry of the MPS can still be partly overcome by increasing the used bond dimension. Quantum chemical DMRG calculations generally need substantially larger bond dimensions than their counterparts in condensed matter physics.

In this dissertation we have introduced the three-legged tree tensor network state (T3NS), a new variational tensor network ansatz for quantum chemical systems. The T3NS is a subclass of the tree tensor network states and pro-

vides an improved flexibility for entanglement encoding compared to DMRG. The branching character of the T3NS also allows a closer packing of correlated orbitals in the network; hence, a lower bond dimension is generally needed for T3NS calculations compared to QC-DMRG. Two-site optimization (an important feature in quantum chemical tensor network calculations) is generally accompanied by an increased computational complexity in general TTNSs. This is mitigated in the T3NS by interspersing physical and branching tensors in the network. In Chapter 2, we discussed the T3NS algorithm for quantum chemistry and showed some calculations comparing it to DMRG. These initial comparisons indeed showed the competitiveness of the T3NS with DMRG; the initial T3NS implementation was able to tackle large systems such as the $[\text{Cu}_2\text{O}_2]^{2-}$ isomers with bond dimensions about three to four times smaller compared to DMRG for the same accuracy.

Symmetries play a major role in the development of approximate methods in quantum chemistry and physics as they substantially improve the efficiency of these methods; correctly identifying the symmetries in the system sheds more light on the possible dependencies between the parameters of the ansatz. The point-group symmetry, the particle conservation, and the total spin conservation present in molecular systems introduce sparsities in the tensors. In addition, the total spin conservation can result in a data compression by using the Wigner-Eckart theorem on each tensor. Symmetry-induced dependencies between the parameters of the ansatz are exposed by writing each tensor as a reduced tensor independent of the magnetic quantum number of the basis states and a suitable Clebsch-Gordan coefficient. The introduced sparsity and data compression of the tensors can considerably speed up any tensor contraction needed.

Exploiting the total spin conservation is well understood and widely used in QC-DMRG. A spin-adapted T3NS implementation is needed if we want to obtain a faithful comparison between DMRG and T3NS. The initial proof-of-concept implementation of the T3NS leveraged only the particle conservation; much of the programmed routines needed to be completely overhauled to accommodate the exploitation of multiplicity-free non-abelian symmetries such as the total spin conservation. The resulting implementation can be found on github [156]. In Chapter 3, we explained how such symmetries can be used in the T3NS. We also showed some numerical results exemplifying the increased efficiency when exploiting these symmetries.

Seniority restricted methods have also enjoyed recent success in the simulation of strongly correlated molecular systems. The seniority number of a wave function is defined as the number of unpaired electrons. For example, a wave function with seniority zero has only empty or doubly occupied orbitals.

Unlike the total spin of the electrons, the seniority number is not a good quantum number of the system. Nevertheless, doubly occupied configuration interaction (DOCI, i.e. seniority zero) solutions seem to capture the majority of the strong correlations in a molecule. This solution can be further improved by including configurations with only a few broken pairs. However, restricting the allowed seniority numbers does not solve the exponential growth of the Hilbert space; an exact treatment of seniority-restricted subspaces quickly becomes unfeasible with increasing system size. By using the seniority number in conjunction with tensor network techniques, approximate calculations can be performed in the exponentially scaling seniority-restricted subspaces at a polynomial cost. In Chapter 4 we applied the same techniques used in Chapter 3 to perform such seniority-restricted T3NS calculations. At low seniority number, these calculations are faster than regular QC-T3NS calculations; especially calculations in the seniority-zero subspace prove to be quite fast as the chemical Hamiltonian projected on the seniority-zero subspace is of lower complexity than the full Hamiltonian. This is illustrated by calculating the DOCI ground state energy in a rubidium phenanthrene coordination complex in a minimal basis set with all electrons correlated. This active space is out of reach for modern full QC-TNS calculations as it consists out of 162 electron pairs and 261 spatial orbitals.

The seniority-restricted T3NS can be used for the systematic study of certain known shortcomings and artifacts of ground state solutions in the DOCI subspace. We use the seniority-restricted T3NS for the study of the dissociation of the nitrogen dimer and the weakly bonding neon dimer. The artificial point-group symmetry breaking of the DOCI ground state in benzene is also studied.

In contrast to condensed matter physics, the mapping of the physical degrees of freedom onto the network is generally not straightforward. Hence, multiple heuristics and cost functions for good orderings have been developed over the years. Some of these methods are based on the structure of the quantum chemical Hamiltonian, some are based on measures existing in quantum information theory. In Chapter 5 we give a brief overview of some of the most popular methods for orbital ordering and optimization. Instead of optimizing the ordering of the orbitals globally, one can also reorder the orbitals locally on an already existing T3NS wave function. By locally swapping the physical degrees of freedom, we can minimize the entanglement in the tensor network. We show some preliminary results of the improvements achieved in accuracy with this optimization.

Future outlook

Although the presented calculations demonstrate the competitiveness of the T3NS with DMRG, it is not yet clear if the T3NS can provide substantial advantages for general quantum chemical calculations. Just as DMRG, the T3NS provides a computationally tractable way of calculating ground state wave functions in quantum chemistry and symmetries are easily exploited. T3NS requires demonstrably lower bond dimensions than the MPS due to an improved entanglement encoding and the better grouping of highly entangled orbitals. However, this does not necessarily result into more efficient calculations regarding CPU time; the favorable scaling of DMRG with respect to increasing bond dimension means that we generally experienced similar accuracies in less computational time compared to the T3NS.

To assess if the T3NS is generally advantageous with respect to DMRG we need to leverage its strong points as much as possible. The increased flexibility of the T3NS with respect to the MPS introduces extra opportunities for the optimization of the orbital mapping since the geometry of the network itself can be changed. Optimizing the geometry of the T3NS can be done in a similar fashion as the optimization of the orbital ordering; by sweeping through the network local permutations of the virtual degrees of freedom allows moving branches around in the network. Although this is not implemented yet, it is currently ongoing research. In the near future, this functionality will hopefully shed some light on the performance of T3NS with respect to DMRG in a given orbital set.

Another option is to further investigate the effect of orbital choice within the T3NS framework. For example, the usage of canonical transformations for lowering the entanglement in the network can be studied for T3NS as previously done for the MPS by Krumnow *et al.* [189]. It is very much possible that the impact of such optimizations will be different for the T3NS than for the MPS due to its different network geometry.

Finally, the current implementation is primarily a proof-of-concept. Considerable efforts have been invested in making an efficient and parallel code, but there is still quite some room for improvement. While the code utilizes the OpenMP API for shared memory parallelization, we could further improve by using distributed memory parallelization on top. Furthermore, the memory requirements of QC-DMRG and QC-T3NS are both dominated by the storage of renormalized operators. As T3NS generally needs a lower bond dimension than DMRG for similar accuracies, QC-T3NS tends to be less strenuous on memory requirements. This could make T3NS particularly suitable for a hybrid CPU-GPU approach in the future.

All this could push T3NS to even increasingly larger system sizes. An as-

assessment of a possible cross-over point between the performance of QC-DMRG and QC-T3NS for general molecules will probably require one or more of the above mentioned improvements.



A

Two-site optimization

A.1. The approximation theorem

Being the workhorse of tensor network methods, it is more than appropriate to at least assign an appendix to the singular value decomposition. Probably the most fundamental theorem of the singular value decomposition is given by the *approximation theorem*, i.e.

Theorem A.1 (Approximation theorem) *Approximating the matrix \mathbf{A} by a matrix \mathbf{B}_k with rank k has an error lower-bounded by*

$$\|\mathbf{A} - \mathbf{B}_k\|_F^2 \geq \sum_{i=k+1}^{\min(n,m)} \sigma_i^2, \quad (\text{A.1})$$

where σ_i are the singular values of \mathbf{A} in descending order and both \mathbf{A} and \mathbf{B}_k are $m \times n$ matrices.

Corollary A.1.1 *The minimum of Eq. (A.1) is obtained for the truncated singular value decomposition of \mathbf{A} , i.e.*

$$\mathbf{B}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^\dagger \quad (\text{A.2})$$

where

$$\mathbf{A} = \sum_{i=1}^{\min(m,n)} \sigma_i \mathbf{u}_i \mathbf{v}_i^\dagger \quad (\text{A.3})$$

is the singular value decomposition of \mathbf{A} .

This theorem is quite often called the Eckart-Young theorem or the Eckart-Young-Mirsky theorem. Although Carl Eckart and Gale Young did rediscover the approximation theorem [280], it was, according to Stewart [281], Erhard Schmidt who first formulated it [16].

Leon Mirsky in his turn extended this theorem from the Frobenius norm to every unitarily invariant norm [282]. A unitarily invariant norm is any norm respecting $\|\mathbf{U}^\dagger \mathbf{A} \mathbf{V}\|_U = \|\mathbf{A}\|_U$ for all unitary matrices \mathbf{U} and \mathbf{V} . His extension of the approximation theorem can be formulated as follows [282]:

Theorem A.2 (Mirsky) *Let ρ_i and σ_i ($i = 1, 2, \dots, n$) be the singular values in descending order of the complex matrices \mathbf{A} and \mathbf{B} respectively. Then*

$$\|\mathbf{A} - \mathbf{B}\| \geq \|\text{diag}(\rho_1 - \sigma_1, \dots, \rho_n - \sigma_n)\| \quad (\text{A.4})$$

for any unitarily invariant norm.

It is quite straightforward to prove Theorem A.1 using this theorem. We choose the Frobenius norm as unitarily invariant norm and \mathbf{B} to be a k -rank matrix in Eq. (A.4). Indeed, we obtain the approximation theorem:

$$\|\mathbf{A} - \mathbf{B}\|_F \geq \sum_{i=1}^k (\rho_i - \sigma_i)^2 + \sum_{i=k+1}^n \rho_i^2 \quad (\text{A.5})$$

$$\|\mathbf{A} - \mathbf{B}\|_F \geq \sum_{i=k+1}^n \rho_i^2. \quad (\text{A.6})$$

When performing a truncated singular value decomposition during a tensor network optimization we are interested in the best rank k approximation to a certain matrix \mathbf{A} under the constraint that the approximation has a fixed norm $\|\mathbf{B}_k\| = 1$; we have to minimize the following Lagrangian:

$$\mathcal{L} = \|\mathbf{A} - \mathbf{B}_k\|^2 - \lambda (\|\mathbf{B}_k\|^2 - 1) \quad (\text{A.7})$$

To solve this, Mirsky's theorem is more useful than the approximation theorem. Indeed, we can state the following corollary.

Corollary A.2.1 *Given the matrix \mathbf{A} and a unitarily invariant norm $\|\cdot\|$. Solving the constrained minimization of $\|\mathbf{A} - \mathbf{B}\|$ for constraints on \mathbf{B} that can be written purely in function of the singular values of \mathbf{B} boils down to solving the minimization problem of $\|\text{diag}(\rho_1 - \sigma_1, \dots, \rho_n - \sigma_n)\|$ under said constraints where*

$$\mathbf{A} = \mathbf{U} \text{diag}(\rho_1, \dots, \rho_n) \mathbf{V}^\dagger$$

is the singular value decomposition of \mathbf{A} . An optimal \mathbf{B} is then given by

$$\mathbf{B} = \mathbf{U} \text{diag}(\sigma_1^*, \dots, \sigma_n^*) \mathbf{V}^\dagger.$$

The set of singular values σ_i^* is a solution for the constrained minimization of $\|\text{diag}(\rho_1 - \sigma_1, \dots, \rho_n - \sigma_n)\|$.

This follows quite naturally from Theorem A.2. Indeed, for any matrix \mathbf{B}' fulfilling the constraints we can construct the matrix \mathbf{B} with the same singular values as \mathbf{B}' but the unitary matrices for its SVD shared with \mathbf{A} . This matrix \mathbf{B} still fulfills the constraints and we have $\|\mathbf{A} - \mathbf{B}'\| \geq \|\mathbf{A} - \mathbf{B}\|$, i.e. \mathbf{B} is equal to or better than \mathbf{B}' for the constrained minimization problem.

Let us prove that, using the Frobenius norm, the best normed rank- k matrix approximating \mathbf{A} indeed corresponds with an ordinary truncated singular value decomposition and renorming the result. From the previous corollary, it is clear that minimizing the Lagrangian in Eq. (A.7) corresponds to minimizing

$$\mathcal{L}' = \|\text{diag}(\rho_1 - \sigma_1, \dots, \rho_k - \sigma_k, \rho_{k+1}, \dots, \rho_n)\|_F^2 - \lambda (\|\text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0)\|_F^2 - 1) \quad (\text{A.8})$$

or

$$\mathcal{L}' = \sum_{i=1}^k (\rho_i - \sigma_i)^2 + \sum_{i=k+1}^n \rho_i^2 - \lambda \left(\sum_{i=1}^k \sigma_i^2 - 1 \right). \quad (\text{A.9})$$

Minimizing this Lagrangian corresponds to solving

$$\frac{\partial \mathcal{L}'}{\partial \sigma_i} = -2(\rho_i - \sigma_i) - 2\lambda \sigma_i = 0 \quad \forall i \quad (\text{A.10})$$

or

$$\sigma_i = \frac{\rho_i}{1 - \lambda}. \quad \forall i \quad (\text{A.11})$$

The rescaling factor can be easily computed as we imposed $\sum_i \sigma_i^2 = 1$.

A very good historical overview of the singular value decomposition is given in ref. [281] together with both Schmidt's and Weyl's proof for the approximation theorem for the Frobenius and spectral norm.

A.2. One-site versus two-site optimization

One-site and two-site optimizations do not necessarily result in the same wave function. In fact, the next example will illustrate that they can be vastly different. Let us investigate a system consisting out of two interacting spin-1/2 particles. The Hamiltonian we are going to study is given by

$$\hat{H} = \mathbf{S}_1 \cdot \mathbf{S}_2 + 2S_1^z + \frac{3}{5}S_2^z, \quad (\text{A.12})$$

i.e. they are subject to a Heisenberg nearest neighbor interaction and a magnetic field in the z-direction is present with a varying strength. Assume that we, for some reason, want to find the ground state of this wave function through an MPS with a bond dimension of 1.

Let us first examine for two-site optimization. Of course the algorithm will find the exact ground state during the optimization, which is given by

$$|\Psi\rangle = 0.9523|\downarrow\uparrow\rangle - 0.3052|\uparrow\downarrow\rangle, \quad (\text{A.13})$$

and has an energy $E = -1.1102$. In the next step of the optimization this two-site optimized wave function will be decomposed into two separate MPSs with a bond dimension of 1. This will result in $|\varphi_{\text{twosite}}\rangle = |\downarrow\uparrow\rangle$ and an energy of $E = -0.95$.

However, when doing one-site optimization, the best mean-field solution for the Hamiltonian can be found, i.e.

$$|\varphi_{\text{onesite}}\rangle = [0.9948|\downarrow\rangle - 0.1021|\uparrow\rangle] \otimes [0.9317|\downarrow\rangle + 0.3633|\uparrow\rangle]. \quad (\text{A.14})$$

This state has an energy of $E = -1.0542$ which is better than the truncated solution from the two-site optimization. On the other hand, its fidelity with respect to the true ground state is much lower:

$$|\langle\varphi_{\text{onesite}}|\Psi\rangle|^2 = 0.1392 \quad (\text{A.15})$$

$$|\langle\varphi_{\text{twosite}}|\Psi\rangle|^2 = 0.9069. \quad (\text{A.16})$$

Of course, we presented here quite an edge-case. The Hamiltonian was tailored to accentuate the differences between the two optimizations and the approximation made by the MPS is not the least; only a bond dimension of one was allowed. When working with MPSs with larger, realistic bond dimensions, the differences between the two optimizations are not expected to be this vast. This example was merely to illustrate the possible different outcomes for the optimizations.

B

Wigner symbols

In this appendix, some of the frequently used operations for the Wigner symbols are given. We will extensively use a graphical representation of the Wigner $3j$ -symbol as introduced in Chapter 3. Here, we will only consider the $SU(2)$ -symmetry; Fermionic signs will be completely disregarded. Results were taken from appendix B of Ref. [62] and from Ref. [283]; they are translated to the graphical representation used in this dissertation.

This said, let us reiterate the relevant graphical depiction limited to the $SU(2)$ -symmetry:

$$\begin{array}{c} \uparrow \mathbf{x}_2 \\ \curvearrowright \\ \leftarrow \mathbf{x}_1 \quad \rightarrow \mathbf{x}_3 \end{array} = (-1)^{j_3 - m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}. \quad (\text{B.1})$$

The arching arrow indicates the order of the coupling within the $3j$ -symbol while the arrows on the edges indicate the sign of their respective spin projections in the $3j$ -symbol. Every outgoing edge of a vertex also introduces a phase $(-1)^{j_i - m_i}$. For example, the following graphical depiction corresponds with

$$\begin{array}{c} \uparrow \mathbf{x}_3 \\ \leftarrow \mathbf{x}_1 \quad \rightarrow \mathbf{x}_2 \end{array} = (-1)^{j_1 - m_1} (-1)^{j_3 - m_3} \begin{pmatrix} j_1 & j_3 & j_2 \\ -m_1 & -m_3 & m_2 \end{pmatrix}. \quad (\text{B.2})$$

B.1. The Wigner 3j-symbol

The relation between the Wigner 3j-symbols and the Clebsch-Gordan coefficients is given by

$$\langle j_1 m_1 j_2 m_2 | j_3 m_3 \rangle = \sqrt{2j_3 + 1} (-1)^{j_1 - j_2 + m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \quad (\text{B.3})$$

$$= (-1)^{j_2 - j_1 - j_3} [j_3] \begin{array}{c} \downarrow \mathbf{x}_2 \\ \circlearrowleft \\ \mathbf{x}_1 \rightarrow \mathbf{x}_3 \end{array}, \quad (\text{B.4})$$

where $[j_3]$ is shorthand for $\sqrt{2j_3 + 1}$. Both the Wigner 3j-symbols and the Clebsch-Gordan coefficients dictate how to add angular momenta; the 3j-symbols do it more symmetrically, as such, they are preferred for our purposes. The 3j-symbols are invariant for even permutations of the columns while an extra sign of $(-1)^{j_1 + j_2 + j_3}$ is introduced for odd permutations. Even permutations correspond with a rotation of the arching arrow and odd permutations correspond with changing the direction of the arrow, i.e.

$$\begin{array}{c} \mathbf{x}_2 \\ \circlearrowleft \\ \mathbf{x}_1 \rightarrow \mathbf{x}_3 \end{array} = \begin{array}{c} \mathbf{x}_2 \\ \circlearrowright \\ \mathbf{x}_1 \rightarrow \mathbf{x}_3 \end{array} = \begin{array}{c} \mathbf{x}_2 \\ \circlearrowleft \\ \mathbf{x}_1 \leftarrow \mathbf{x}_3 \end{array} \quad (\text{rotation})$$

and

$$\begin{array}{c} \mathbf{x}_2 \\ \circlearrowleft \\ \mathbf{x}_1 \rightarrow \mathbf{x}_3 \end{array} = (-1)^{j_1 + j_2 + j_3} \begin{array}{c} \mathbf{x}_2 \\ \circlearrowright \\ \mathbf{x}_1 \rightarrow \mathbf{x}_3 \end{array} \quad (\text{reflection})$$

irrespective of the directions of the edges (signs of m).

Orthogonality for the 3j-symbols is given by

$$\sum_{m_1 m_2} [j_3]^2 \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j'_3 \\ m_1 & m_2 & m'_3 \end{pmatrix} = \delta_{j_3, j'_3} \delta_{m_3, m'_3} \quad (\text{B.5})$$

or

$$\sum_{m_1 m_2} [j_3]^2 \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_2 & j_1 & j'_3 \\ -m_2 & -m_1 & -m'_3 \end{pmatrix} = \delta_{j_3, j'_3} \delta_{m_3, m'_3} \quad (\text{B.6})$$

or

$$\begin{array}{c} \mathbf{x}_1 \\ \circlearrowleft \\ \mathbf{x}_3 \leftarrow \mathbf{x}_2 \end{array} = (-1)^{j_1 + j_2 - j_3} [j_3]^{-2} \begin{array}{c} \mathbf{x}_1 \\ \circlearrowright \\ \mathbf{x}_3 \rightarrow \mathbf{x}_2 \end{array}. \quad (\text{orthogonality})$$

An implicit summation over m_1 and m_2 is assumed.

Last, we can change the direction of an internal edge; as an internal contraction implies a summation over its corresponding m_i we can just change to $-m_i$. The only thing that should be taken into account is that the phase

$(-1)^{j_i - m_i}$ of the internal edge changes to $(-1)^{2j_i}(-1)^{j_i - m_i}$. Reversal is thus given by

$$\begin{array}{c} \rightarrow \\ \text{---} \text{---} \text{---} \\ \leftarrow \\ \mathbf{x}_i \end{array} = (-1)^{2j_i} \begin{array}{c} \rightarrow \\ \text{---} \text{---} \text{---} \\ \leftarrow \\ \mathbf{x}_i \end{array}. \quad (\text{reversal})$$

Just as the Clebsch-Gordan coefficients, the Wigner $3j$ -symbols are trivially zero except for valid couplings, i.e. for a $3j$ -symbol $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ we have

$$m_i \in \{-j_i, -j_i + 1, \dots, j_i - 1, j_i\} \quad (\text{B.7})$$

$$m_1 + m_2 + m_3 = 0 \quad (\text{B.8})$$

$$j_3 \in \{|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 - 1, j_1 + j_2\}. \quad (\text{B.9})$$

B.2. The Wigner $6j$ -symbol

When coupling three different angular momenta to a total angular momentum, there are three different ways to achieve this. Depending on which spins are first coupled to an intermediate total spin, we obtain different coupling schemes. Wigner $6j$ -symbols define the connection between these different coupling schemes, i.e.

$$\begin{aligned} |(j_1, (j_2, j_3)J_{23})jm) &= \sum_{J_{12}} (-1)^{j_1 + j_2 + j_3 + J_{12}} [J_{12}] [J_{23}] \\ &\times \begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & j & J_{23} \end{Bmatrix} |((j_1, j_2)J_{12}, j_3)jm). \end{aligned} \quad (\text{B.10})$$

The coupling order are given in between brackets. We will use the Wigner $6j$ -symbol extensively for the following symmetry tensor operation:

$$\begin{array}{c} j_1 \\ \diagdown \quad \diagup \\ \circ \quad \circ \\ \diagup \quad \diagdown \\ j_2 \quad j_4 \\ \leftarrow \quad \rightarrow \\ j \end{array} = \sum_{j'} [j']^2 (-1)^{j + j' - j_1 - j_3} \begin{Bmatrix} j_1 & j_4 & j' \\ j_3 & j_2 & j \end{Bmatrix} \begin{array}{c} j_3 \\ \diagdown \quad \diagup \\ \circ \quad \circ \\ \diagup \quad \diagdown \\ j_2 \quad j_4 \\ \leftarrow \quad \rightarrow \\ j' \end{array} \quad (6j\text{-swap})$$

In this equation, we have swapped the angular momenta j_1 and j_3 by using a $6j$ -symbol. This operation is independent of the directions of the arrows on the loose edges.

B.3. The Wigner $9j$ -symbol

While the Wigner $6j$ -symbol was used to change the coupling scheme of three angular momenta, the Wigner $9j$ -symbol is used when recoupling four angu-

C.3. Some more contractions

C.3.1. Physical update of a renormalized operator

In Section 3.4.2 the construction of an extended renormalized operator from a minimal renormalized operator by appending a site operator was exemplified. Here, we will show how said extended renormalized operator \mathbf{R} can again be updated by using an optimized canonicalized tensor \mathbf{B} to a minimal renormalized operator \mathbf{S} :

C

$$\begin{aligned}
 & \begin{array}{ccc}
 x_H \text{ --- } S^r & & x_H \text{ --- } \text{---} \\
 \begin{array}{c} | \\ x_{2'}\alpha_{2'} \\ | \\ x_2\alpha_2 \end{array} & & \begin{array}{c} \uparrow \\ x_{2'}\alpha_{2'} \\ \uparrow \\ x_2\alpha_2 \end{array} \\
 x_{2'}\alpha_{2'} \text{ --- } B^r & & x_{2'}\alpha_{2'} \text{ --- } \text{---} \\
 \begin{array}{c} \curvearrowright \\ x_{p'} \\ \curvearrowleft \end{array} & & \begin{array}{c} \curvearrowright \\ x_{p'} \\ \curvearrowleft \end{array} \\
 x_H \text{ --- } R^r & & x_H \text{ --- } \text{---} \\
 \begin{array}{c} | \\ x_{3'}\alpha_{3'} \\ | \\ x_p \end{array} & & \begin{array}{c} \uparrow \\ x_{3'}\alpha_{3'} \\ \uparrow \\ x_p \end{array} \\
 x_2\alpha_2 \text{ --- } B^r & & x_2\alpha_2 \text{ --- } \text{---} \\
 \begin{array}{c} \curvearrowright \\ x_p \\ \curvearrowleft \end{array} & & \begin{array}{c} \curvearrowright \\ x_p \\ \curvearrowleft \end{array} \\
 x_3\alpha_3 & & x_3\alpha_3
 \end{array}
 \end{aligned}
 \tag{C.11}$$

$= \underbrace{(-1)^{\pi_{p'}} (-1)^{j_{2'} + j_{p'} - j_3'}}_{\text{Due to the adjoint}}$

where an extra phase is included for the adjoint of a suitably orthogonal tensor, see Section 3.2.2. By using the orthogonality and the reversal property twice, we obtain

$$\begin{aligned}
 & \begin{array}{ccc}
 x_H \text{ --- } S^r & = & x_H \text{ --- } R^r \\
 \begin{array}{c} | \\ x_{2'}\alpha_{2'} \\ | \\ x_2\alpha_2 \end{array} & & \begin{array}{c} \curvearrowright \\ x_{p'} \\ \curvearrowleft \end{array} \\
 & & & \begin{array}{c} | \\ x_{3'}\alpha_{3'} \\ | \\ x_p \end{array} \\
 & & & \begin{array}{c} \curvearrowright \\ x_p \\ \curvearrowleft \end{array} \\
 & & & x_2\alpha_2 \text{ --- } B^r \\
 & & & \begin{array}{c} \curvearrowright \\ x_p \\ \curvearrowleft \end{array} \\
 & & & x_3\alpha_3
 \end{array}
 \end{aligned}
 \tag{C.12}$$

$$\delta_{x_2, x_{3p}} \delta_{x_{2'}, x_{3'p'}} \underbrace{(-1)^{\pi_p}}_{\text{Fermionic}} \underbrace{[j_2]^{-2} [j_{2'}]^{-2} (-1)^{j_2 - j_3 + j_p}}_{\text{SU}(2)}$$

C.3.2. Branching update of a renormalized operator

Not only updates through physical tensors are possible, we can also update two minimal renormalized operators to a new minimal renormalized operator by

fusing it together with a branching tensor, i.e.

$$\begin{aligned}
 & \begin{array}{c} x_{3'}\alpha_{3'} \\ | \\ \text{Tr}^r \\ | \\ x_3\alpha_3 \end{array} x_T \quad \begin{array}{c} \downarrow \\ \mathbf{x}_{3'} \\ \downarrow \\ \mathbf{x}_3 \end{array} \quad \begin{array}{c} \downarrow \\ \mathbf{x}_T \end{array} \\
 & \begin{array}{c} x_{2'}\alpha_{2'} \\ | \\ \text{Ar}^r \\ | \\ x_2\alpha_2 \end{array} \quad \begin{array}{c} x_{3'}\alpha_{3'} \\ | \\ \text{Ar}^r \\ | \\ x_3\alpha_3 \end{array} \\
 & \begin{array}{c} x_{1'}\alpha_{1'} \\ | \\ \text{Sr}^r \\ | \\ x_1\alpha_1 \end{array} \quad \begin{array}{c} x_S \\ | \\ \text{C} \\ | \\ x_R \end{array} \quad \begin{array}{c} x_T \end{array} \\
 & = (-1)^{j_{1'}+j_{2'}-j_3} \quad \begin{array}{c} \mathbf{x}_{1'} \\ \downarrow \\ \mathbf{x}_2 \\ \downarrow \\ \mathbf{x}_1 \end{array} \quad \begin{array}{c} \mathbf{x}_{2'} \\ \downarrow \\ \mathbf{x}_3 \\ \downarrow \\ \mathbf{x}_1 \end{array} \quad \begin{array}{c} \mathbf{x}_S \\ \downarrow \\ \mathbf{x}_R \\ \downarrow \\ \mathbf{x}_T \end{array} \quad \text{(C.13)}
 \end{aligned}$$

where C is the needed factor for correctly fusing the two renormalized operators together. By using the 9j-merge, reflection and reversal property we obtain

$$\begin{aligned}
 & \begin{array}{c} x_{3'}\alpha_{3'} \\ | \\ \text{Tr}^r \\ | \\ x_3\alpha_3 \end{array} x_T = \begin{array}{c} x_{2'}\alpha_{2'} \\ | \\ \text{Ar}^r \\ | \\ x_2\alpha_2 \end{array} \quad \begin{array}{c} x_{3'}\alpha_{3'} \\ | \\ \text{Ar}^r \\ | \\ x_3\alpha_3 \end{array} \\
 & \begin{array}{c} x_{1'}\alpha_{1'} \\ | \\ \text{Sr}^r \\ | \\ x_1\alpha_1 \end{array} \quad \begin{array}{c} x_S \\ | \\ \text{C} \\ | \\ x_R \end{array} \quad \begin{array}{c} x_T \end{array} \quad \underbrace{(-1)^{\pi_R \pi_2}}_{\text{Fermionic}} \underbrace{\left\{ \begin{array}{ccc} j_1 & j_{1'} & j_R \\ j_2 & j_{2'} & j_S \\ j_3 & j_{3'} & j_T \end{array} \right\}}_{\text{SU}(2)} \quad \text{(C.14)}
 \end{aligned}$$

C.3.3. The effective Hamiltonian: The DMRG case

In Section 3.4.3 the action of the effective Hamiltonian on a physical and a branching tensor was exemplified. Let us now discuss the action of the effective Hamiltonian on two physical sites instead. The two-site tensor to optimize is now given by

$$\mathbf{M} = \begin{array}{c} x_i \\ | \\ \text{M}_{x_2}^r \\ | \\ x_j \\ | \\ x_{3'}\alpha_3 \end{array} \quad \begin{array}{c} \downarrow \\ \mathbf{x}_i \\ \downarrow \\ \mathbf{x}_j \end{array} \quad \begin{array}{c} \downarrow \\ \mathbf{x}_{3'}\alpha_3 \end{array} \quad \text{(C.15)}$$

Two sets of extended renormalized operators are needed as given in Table 3.1, one of each flavor (upstream and downstream). The action of these renormalized operators upon the two-site tensor is given by

$$\mathbf{H} = \begin{array}{c} x_{i'} \\ | \\ \text{H}_{x_2}^r \\ | \\ x_{j'} \\ | \\ x_{3'}\alpha_{3'} \end{array} \quad \begin{array}{c} \downarrow \\ \mathbf{x}_{i'} \\ \downarrow \\ \mathbf{x}_{j'} \end{array} \quad \begin{array}{c} \downarrow \\ \mathbf{x}_{3'}\alpha_{3'} \end{array} \quad \text{(C.16)}$$

$$\begin{aligned}
 & = \begin{array}{c} x_{1'}\alpha_{1'} \\ | \\ \text{R}_{x_{1'}x_{1'}}^r \\ | \\ x_{i'} \\ | \\ \text{S}_{x_{3'}x_{3'}}^r \\ | \\ x_{j'} \\ | \\ x_{3'}\alpha_{3'} \end{array} \quad \begin{array}{c} x_H \\ | \\ \text{M}_{x_2}^r \\ | \\ x_j \\ | \\ x_3\alpha_3 \end{array} \\
 & \begin{array}{c} \mathbf{x}_{1'i'} \\ \downarrow \\ \mathbf{x}_{1i} \\ \downarrow \\ \mathbf{x}_i \end{array} \quad \begin{array}{c} \mathbf{x}_{3'j'} \\ \downarrow \\ \mathbf{x}_{3j} \\ \downarrow \\ \mathbf{x}_j \end{array} \quad \begin{array}{c} \mathbf{x}_H \\ \downarrow \\ \mathbf{x}_i \\ \downarrow \\ \mathbf{x}_j \end{array} \quad \text{(C.17)}
 \end{aligned}$$

By using three times orthogonality and reversal once we obtain

$$\underbrace{(-1)^{\pi_j}}_{\text{Fermionic}} [j_2]^{-2} [j_{3j}]^{-2} [j_{3'j'}]^{-2} (-1)^{(2j_j) + (j_j + j_3 - j_{3j}) + (j_i + j_1 - j_2) + (j_H + j_{3j} - j_{3'j'})} \underbrace{\hspace{10em}}_{\text{SU}(2)} \tag{C.18}$$

The effective Hamiltonian of DMRG can be done by a tensor contraction without the need of any Wigner symbols.



Bibliography

- [1] P. Ehrenfest, *Ansprache zur Verleihung der Lorentzmedaille an Professor Wolfgang Pauli am 31. Oktober 1931* (Springer, 1988), pp. 43–48.
- [2] A. Comte, *Cours de philosophie positive*, Vol. La Philosophie chimique et la philosophie de la physique, 3 (Bachelier, 1838), pp. 44–45.
- [3] E. Schrödinger, “Quantisierung als Eigenwertproblem”, *Annalen der Physik* **384**, 361–376 (1926).
- [4] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?”, *Physical Review* **47**, 777–780 (1935).
- [5] D. Bohm and Y. Aharonov, “Discussion of experimental proof for the paradox of Einstein, Rosen, and Podolsky”, *Physical Review* **108**, 1070–1076 (1957).
- [6] A. Einstein and M. Born, *Briefwechsel 1916 - 1955* (Langen Müller, München, 2005).
- [7] L. de Broglie, “La mécanique ondulatoire et la structure atomique de la matière et du rayonnement”, *Journal de Physique et le Radium* **8**, 225–241 (1927).
- [8] D. Bohm, “A suggested interpretation of the quantum theory in terms of “hidden” variables. I”, *Physical Review* **85**, 166–179 (1952).
- [9] D. Bohm, “A suggested interpretation of the quantum theory in terms of “hidden” variables. II”, *Physical Review* **85**, 180–193 (1952).
- [10] J. S. Bell, “On the Einstein Podolsky Rosen paradox”, *Physics Physique Fizika* **1**, 195–200 (1964).
- [11] S. J. Freedman and J. F. Clauser, “Experimental test of local hidden-variable theories”, *Physical Review Letters* **28**, 938–941 (1972).
- [12] A. Aspect, P. Grangier, and G. Roger, “Experimental tests of realistic local theories via Bell’s theorem”, *Physical Review Letters* **47**, 460–463 (1981).
- [13] B. Hensen et al., “Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres”, *Nature* **526**, 682–686 (2015).

- [14] M. Giustina et al., “Significant-loophole-free test of Bell’s theorem with entangled photons”, *Physical Review Letters* **115**, 250401 (2015).
- [15] L. K. Shalm et al., “Strong loophole-free test of local realism”, *Physical Review Letters* **115**, 250402 (2015).
- [16] E. Schmidt, “Zur Theorie der linearen und nichtlinearen Integralgleichungen”, in *Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten*, edited by A. Pietsch (Vieweg+Teubner Verlag, Wiesbaden, 1989), pp. 190–233.
- [17] J. von Neumann, “Thermodynamik quantenmechanischer Gesamtheiten”, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 273–291 (1927).
- [18] A. Rényi et al., “On measures of entropy and information”, in *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics* (The Regents of the University of California, 1960), pp. 547–561.
- [19] W. van Dam and P. Hayden, “Renyi-entropic bounds on quantum communication”, arXiv e-prints, quant-ph/0204093 (2002).
- [20] M. B. Hastings, “An area law for one-dimensional quantum systems”, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, P08024–P08024 (2007).
- [21] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer, “Entropy, entanglement, and area: analytical results for harmonic lattice systems”, *Physical Review Letters* **94**, 060503 (2005).
- [22] J. Eisert, M. Cramer, and M. B. Plenio, “Colloquium: area laws for the entanglement entropy”, *Reviews of Modern Physics* **82**, 277–306 (2010).
- [23] K. Van Acoleyen, M. Mariën, and F. Verstraete, “Entanglement rates and area laws”, *Physical Review Letters* **111**, 170501 (2013).
- [24] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, “Entanglement in quantum critical phenomena”, *Physical Review Letters* **90**, 227902 (2003).
- [25] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information* (American Association of Physics Teachers, 2010).
- [26] P. A. M. Dirac and R. H. Fowler, “The quantum theory of the electron”, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **117**, 610–624 (1928).
- [27] P. Pyykko, “Relativistic effects in structural chemistry”, *Chemical Reviews* **88**, 563–594 (1988).

- [28] B. A. Hess, “Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators”, *Physical Review A* **33**, 3742–3748 (1986).
- [29] M. Krauss and W. J. Stevens, “Effective potentials in molecular quantum chemistry”, *Annual Review of Physical Chemistry* **35**, 357–385 (1984).
- [30] R. Mastalerz and M. Reiher, “Relativistic electronic structure theory for molecular spectroscopy”, in *Handbook of high-resolution spectroscopy* (American Cancer Society, 2011).
- [31] M. Reiher, “Relativistic Douglas–Kroll–Hess theory”, *WIREs Computational Molecular Science* **2**, 139–149 (2012).
- [32] P. Tecmer, K. Boguslawski, and D. Kędziera, “Relativistic methods in computational quantum chemistry”, in *Handbook of computational chemistry* (Springer Netherlands, 2016), pp. 1–43.
- [33] T. Saue, “Relativistic Hamiltonians for chemistry: a primer”, *ChemPhysChem* **12**, 3077–3094 (2011).
- [34] D. R. Hartree, “The wave mechanics of an atom with a non-Coulomb central field. part I. theory and methods”, *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 89–110 (1928).
- [35] M. Born and R. Oppenheimer, “Zur Quantentheorie der Molekeln”, *Annalen der Physik* **389**, 457–484 (1927).
- [36] E. A. Hylleraas, “Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium”, *Zeitschrift für Physik* **54**, 347–366 (1929).
- [37] T. Kinoshita, “Ground state of the helium atom”, *Physical Review* **105**, 1490–1502 (1957).
- [38] K. Frankowski and C. L. Pekeris, “Logarithmic terms in the wave functions of the ground state of two-electron atoms”, *Physical Review* **146**, 46–49 (1966).
- [39] C. Schwartz, “Experiment and theory in computations of the He atom ground state”, *International Journal of Modern Physics E* **15**, 877–888 (2006).
- [40] C. Cerjan, ed., *Numerical grid methods and their application to Schrödinger’s equation* (Springer Netherlands, 1993).
- [41] S. K. Gray and E. M. Goldfield, “Dispersion fitted finite difference method with applications to molecular quantum mechanics”, *The Journal of Chemical Physics* **115**, 8331–8344 (2001).

- [42] J. R. Flores, E. Clementi, and V. Sonnad, "Solution of atomic Hartree-Fock equations with the P version of the finite element method", *The Journal of Chemical Physics* **91**, 7030–7038 (1989).
- [43] S. R. White, J. W. Wilkins, and M. P. Teter, "Finite-element method for electronic structure", *Physical Review B* **39**, 5819–5833 (1989).
- [44] E. Tsuchida and M. Tsukada, "Electronic-structure calculations based on the finite-element method", *Physical Review B* **52**, 5573–5578 (1995).
- [45] D. Sundholm, "Finite-element multiconfiguration Hartree-Fock calculations of the atomic quadrupole moment of $\text{Ar}^+(^2P_{3/2})$ ", *Physical Review A* **59**, 3355–3358 (1999).
- [46] R. Alizadegan, K. J. Hsia, and T. J. Martinez, "A divide and conquer real space finite-element Hartree-Fock method", *The Journal of Chemical Physics* **132**, 034101 (2010).
- [47] S. F. Boys and A. C. Egerton, "Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system", *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **200**, 542–554 (1950).
- [48] J. A. Pople and W. J. Hehre, "Computation of electron repulsion integrals involving contracted Gaussian basis functions", *Journal of Computational Physics* **27**, 161–168 (1978).
- [49] W. J. Hehre, R. F. Stewart, and J. A. Pople, "Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals", *The Journal of Chemical Physics* **51**, 2657–2664 (1969).
- [50] J. C. Slater, "Atomic shielding constants", *Physical Review* **36**, 57–64 (1930).
- [51] D. R. Hartree, "The wave mechanics of an atom with a non-Coulomb central field. Part II. Some results and discussion", *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 111–132 (1928).
- [52] J. C. Slater, "The self consistent field and the structure of atoms", *Physical Review* **32**, 339–348 (1928).
- [53] V. Fock, "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems", *Zeitschrift für Physik* **61**, 126–148 (1930).
- [54] J. M. Foster and S. F. Boys, "Canonical configurational interaction procedure", *Reviews of Modern Physics* **32**, 300–302 (1960).

- [55] C. Edmiston and K. Ruedenberg, “Localized atomic and molecular orbitals”, *Reviews of Modern Physics* **35**, 457–464 (1963).
- [56] J. Pipek and P. G. Mezey, “A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions”, *The Journal of Chemical Physics* **90**, 4916–4926 (1989).
- [57] P. A. M. Dirac and N. H. D. Bohr, “The quantum theory of the emission and absorption of radiation”, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **114**, 243–265 (1927).
- [58] V. Fock, “Konfigurationsraum und zweite Quantelung”, *Zeitschrift für Physik* **75**, 622–647 (1932).
- [59] W. Pauli, “Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren”, *Zeitschrift für Physik* **31**, 765–783 (1925).
- [60] N. Straumann, “The role of the exclusion principle for atoms to stars: a historical account”, arXiv e-prints, quant-ph/0403199 (2004).
- [61] J. C. Slater, “The theory of complex spectra”, *Physical Review* **34**, 1293–1322 (1929).
- [62] W. H. Dickhoff and D. Van Neck, *Many-body theory exposed!* (World Scientific, Apr. 2005).
- [63] J. Cornwell, *Group theory in physics* (Elsevier, 1997).
- [64] D. R. Herrick, “New symmetry properties of atoms and molecules”, in *Advances in chemical physics* (John Wiley & Sons, Ltd, 2007), pp. 1–115.
- [65] M. Tinkham and G. McKay, *Group theory and quantum mechanics* (Dover, 2003).
- [66] V. Fock, “Zur Theorie des Wasserstoffatoms”, *Zeitschrift für Physik* **98**, 145–154 (1935).
- [67] H. Weyl, “Quantenmechanik und Gruppentheorie”, *Zeitschrift für Physik* **46**, 1–46 (1927).
- [68] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas”, *Physical Review* **136**, B864–B871 (1964).
- [69] W. Kohn and L. J. Sham, “Self-consistent equations including exchange and correlation effects”, *Physical Review* **140**, A1133–A1138 (1965).

- [70] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko, “Density functional theory is straying from the path toward the exact functional”, *Science* **355**, 49–52 (2017).
- [71] J. C. Slater, “Note on Hartree’s method”, *Physical Review* **35**, 210–211 (1930).
- [72] D. R. Hartree and W. Hartree, “Self-consistent field, with exchange, for beryllium”, *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences* **150**, 9–33 (1935).
- [73] J. A. Pople and R. K. Nesbet, “Self-consistent orbitals for radicals”, *The Journal of Chemical Physics* **22**, 571–572 (1954).
- [74] P.-O. Löwdin, “Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects”, *Physical Review* **97**, 1509–1520 (1955).
- [75] T. Helgaker, P. Jørgensen, and J. Olsen, “The standard models”, in *Molecular electronic-structure theory* (John Wiley & Sons, Ltd, 2014) Chap. 5, pp. 142–200.
- [76] R. J. Bartlett and M. Musiał, “Coupled-cluster theory in quantum chemistry”, *Reviews of Modern Physics* **79**, 291–352 (2007).
- [77] R. J. Bartlett and J. F. Stanton, “Applications of post-Hartree-Fock methods: a tutorial”, in *Reviews in computational chemistry* (John Wiley & Sons, Inc., Jan. 2007), pp. 65–169.
- [78] O. Sinanoğlu and D. F.-t. Tuan, “Many-electron theory of atoms and molecules. III. Effect of correlation on orbitals”, *The Journal of Chemical Physics* **38**, 1740–1748 (1963).
- [79] C. Møller and M. S. Plesset, “Note on an approximation treatment for many-electron systems”, *Physical Review* **46**, 618–622 (1934).
- [80] E. U. Condon, “The theory of complex spectra”, *Physical Review* **36**, 1121–1133 (1930).
- [81] F. Coester, “Bound states of a many-particle system”, *Nuclear Physics* **7**, 421–424 (1958).
- [82] F. Coester and H. Kümmel, “Short-range correlations in nuclear wave functions”, *Nuclear Physics* **17**, 477–485 (1960).
- [83] Čížek, Jiří, “On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in urself-type expansion using quantum-field theoretical methods”, *The Journal of Chemical Physics* **45**, 4256–4266 (1966).

- [84] B. O. Roos, P. R. Taylor, and P. E. Siegbahn, "A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach", *Chemical Physics* **48**, 157–173 (1980).
- [85] B. O. Roos, "The complete active space SCF method in a fock-matrix-based super-CI formulation", *International Journal of Quantum Chemistry* **18**, 175–189 (1980).
- [86] P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos, "The complete active space SCF (CASSCF) method in a Newton–Raphson formulation with application to the HNO molecule", *The Journal of Chemical Physics* **74**, 2384–2396 (1981).
- [87] S. R. White, "Density matrix formulation for quantum renormalization groups", *Physical Review Letters* **69**, 2863–2866 (1992).
- [88] S. R. White, "Density-matrix algorithms for quantum renormalization groups", *Physical Review B* **48**, 10345–10356 (1993).
- [89] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, "Second-order perturbation theory with a CASSCF reference function", *The Journal of Physical Chemistry* **94**, 5483–5488 (1990).
- [90] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, "Second-order perturbation theory with a complete active space self-consistent field reference function", *The Journal of Chemical Physics* **96**, 1218–1226 (1992).
- [91] Y. Kurashige and T. Yanai, "Second-order perturbation theory with a density matrix renormalization group self-consistent field reference function: theory and application to the study of chromium dimer", *The Journal of Chemical Physics* **135**, 094104 (2011).
- [92] S. Wouters, V. Van Speybroeck, and D. Van Neck, "DMRG-CASPT2 study of the longitudinal static second hyperpolarizability of all-trans polyenes", *The Journal of Chemical Physics* **145**, 054120 (2016).
- [93] S. Sharma and G. K.-L. Chan, "Communication: a flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states", *The Journal of Chemical Physics* **141**, 111101 (2014).
- [94] S. Guo, Z. Li, and G. K.-L. Chan, "A perturbative density matrix renormalization group algorithm for large active spaces", *Journal of Chemical Theory and Computation* **14**, 4063–4071 (2018).
- [95] M. Musiał, A. Perera, and R. J. Bartlett, "Multireference coupled-cluster theory: the easy way", *The Journal of Chemical Physics* **134**, 114108 (2011).

- [96] A. Köhn, M. Hanauer, L. A. Mück, T.-C. Jagau, and J. Gauss, “State-specific multireference coupled-cluster theory”, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **3**, 176–197 (2012).
- [97] F. A. Evangelista, “Perspective: multireference coupled cluster theories of dynamical electron correlation”, *The Journal of Chemical Physics* **149**, 030901 (2018).
- [98] O. Hino, T. Kinoshita, G. K.-L. Chan, and R. J. Bartlett, “Tailored coupled cluster singles and doubles method applied to calculations on molecular structure and harmonic vibrational frequencies of ozone”, *The Journal of Chemical Physics* **124**, 114311 (2006).
- [99] L. Veis, A. Antalík, J. Brabec, F. Neese, Ö. Legeza, and J. Pittner, “Coupled cluster method with single and double excitations tailored by matrix product state wave functions”, *The Journal of Physical Chemistry Letters* **7**, 4072–4078 (2016).
- [100] F. M. Faulstich et al., “Numerical and theoretical aspects of the DMRG-TCC method exemplified by the nitrogen dimer”, *Journal of Chemical Theory and Computation* **15**, 2206–2220 (2019).
- [101] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley & Sons, Ltd, 2014).
- [102] P. Jørgensen, *Second quantization-based methods in quantum chemistry* (Elsevier, 1981).
- [103] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Courier Corporation, 2012).
- [104] K. Gunst, F. Verstraete, S. Wouters, Ö. Legeza, and D. Van Neck, “T3NS: three-legged tree tensor network states”, *Journal of Chemical Theory and Computation* **14**, 2026–2033 (2018).
- [105] K. Gunst, F. Verstraete, and D. Van Neck, “Three-legged tree tensor networks with SU(2) and molecular point group symmetry”, *Journal of Chemical Theory and Computation* **15**, 2996–3007 (2019).
- [106] S. Tomonaga, “On a relativistically invariant formulation of the quantum theory of wave fields*”, *Progress of Theoretical Physics* **1**, 27–42 (1946).
- [107] J. Schwinger, “Quantum electrodynamics. I. A covariant formulation”, *Physical Review* **74**, 1439–1461 (1948).
- [108] J. Schwinger, “On quantum-electrodynamics and the magnetic moment of the electron”, *Physical Review* **73**, 416–417 (1948).

- [109] R. P. Feynman, “Space-time approach to quantum electrodynamics”, *Physical Review* **76**, 769–789 (1949).
- [110] R. P. Feynman, “The theory of positrons”, *Physical Review* **76**, 749–759 (1949).
- [111] R. P. Feynman, “Mathematical formulation of the quantum theory of electromagnetic interaction”, *Physical Review* **80**, 440–457 (1950).
- [112] F. J. Dyson, “The radiation theories of Tomonaga, Schwinger, and Feynman”, *Physical Review* **75**, 486–502 (1949).
- [113] F. J. Dyson, “The S matrix in quantum electrodynamics”, *Physical Review* **75**, 1736–1755 (1949).
- [114] D. Hanneke, S. F. Hoogerheide, and G. Gabrielse, “Cavity control of a single-electron quantum cyclotron: measuring the electron magnetic moment”, *Physical Review A* **83**, 052122 (2011).
- [115] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, “Tenth-order QED contribution to the electron $g - 2$ and an improved value of the fine structure constant”, *Physical Review Letters* **109**, 111807 (2012).
- [116] K. G. Wilson, “Renormalization group and critical phenomena. II. Phase-space cell analysis of critical behavior”, *Physical Review B* **4**, 3184–3205 (1971).
- [117] K. G. Wilson, “Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture”, *Physical Review B* **4**, 3174–3183 (1971).
- [118] K. G. Wilson and M. E. Fisher, “Critical exponents in 3.99 dimensions”, *Physical Review Letters* **28**, 240–243 (1972).
- [119] K. G. Wilson, “The renormalization group: critical phenomena and the Kondo problem”, *Reviews of Modern Physics* **47**, 773–840 (1975).
- [120] K. G. Wilson and J. Kogut, “The renormalization group and the ϵ expansion”, *Physics Reports* **12**, 75–199 (1974).
- [121] S. R. White and R. M. Noack, “Real-space quantum renormalization groups”, *Physical Review Letters* **68**, 3487–3490 (1992).
- [122] T. Xiang, “Density-matrix renormalization-group method in momentum space”, *Physical Review B* **53**, R10445–R10448 (1996).
- [123] S. R. White and R. L. Martin, “Ab initio quantum chemistry using the density matrix renormalization group”, *The Journal of Chemical Physics* **110**, 4127–4130 (1999).

- [124] 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**, 457–462 (1998).
- [125] S. Östlund and S. Rommer, “Thermodynamic limit of density matrix renormalization”, *Physical Review Letters* **75**, 3537–3540 (1995).
- [126] F. Verstraete, D. Porras, and J. I. Cirac, “Density matrix renormalization group and periodic boundary conditions: a quantum information perspective”, *Physical Review Letters* **93**, 227205 (2004).
- [127] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, “Variational optimization algorithms for uniform matrix product states”, *Physical Review B* **97**, 045145 (2018).
- [128] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, “Time-dependent variational principle for quantum lattices”, *Physical Review Letters* **107**, 070601 (2011).
- [129] I. V. Oseledets, “Tensor-train decomposition”, *SIAM Journal on Scientific Computing* **33**, 2295–2317 (2011).
- [130] G. M. Crosswhite, A. C. Doherty, and G. Vidal, “Applying matrix product operators to model systems with long-range interactions”, *Physical Review B* **78**, 035116 (2008).
- [131] F. Fröwis, V. Nebendahl, and W. Dür, “Tensor operators: constructions and applications for long-range interaction systems”, *Physical Review A* **81**, 062337 (2010).
- [132] U. Schollwöck, “The density-matrix renormalization group in the age of matrix product states”, *Annals of Physics* **326**, 96–192 (2011).
- [133] S. Keller, M. Dolfi, M. Troyer, and M. Reiher, “An efficient matrix product operator representation of the quantum chemical Hamiltonian”, *The Journal of Chemical Physics* **143**, 244118, 244118 (2015).
- [134] G. K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, and S. R. White, “Matrix product operators, matrix product states, and ab initio density matrix renormalization group algorithms”, *The Journal of Chemical Physics* **145**, 014102 (2016).
- [135] C. Hubig, I. P. McCulloch, and U. Schollwöck, “Generic construction of efficient matrix product operators”, *Physical Review B* **95**, 035129 (2017).

- [136] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, “Matrix product density operators: simulation of finite-temperature and dissipative systems”, *Physical Review Letters* **93**, 207204 (2004).
- [137] J. Haegeman and F. Verstraete, “Diagonalizing transfer matrices and matrix product operators: a medley of exact and computational methods”, *Annual Review of Condensed Matter Physics* **8**, 355–406 (2017).
- [138] S. R. White, “Density matrix renormalization group algorithms with a single center site”, *Physical Review B* **72**, 180403 (2005).
- [139] D. Zgid and M. Nooijen, “Obtaining the two-body density matrix in the density matrix renormalization group method”, *The Journal of Chemical Physics* **128**, 144115 (2008).
- [140] C. Hubig, I. P. McCulloch, U. Schollwöck, and F. A. Wolf, “Strictly single-site DMRG algorithm with subspace expansion”, *Physical Review B* **91**, 155115 (2015).
- [141] M. Roemelt and D. A. Pantazis, “Multireference approaches to spin-state energetics of transition metal complexes utilizing the density matrix renormalization group”, *Advanced Theory and Simulations* **2**, 1800201 (2019).
- [142] Q. M. Phung, S. Wouters, and K. Pierloot, “Cumulant approximated second-order perturbation theory based on the density matrix renormalization group for transition metal complexes: a benchmark study”, *Journal of Chemical Theory and Computation* **12**, 4352–4361 (2016).
- [143] T. V. Harris, Y. Kurashige, T. Yanai, and K. Morokuma, “Ab initio density matrix renormalization group study of magnetic coupling in dinuclear iron and chromium complexes”, *The Journal of Chemical Physics* **140**, 054303 (2014).
- [144] S. Wouters, T. Bogaerts, P. Van Der Voort, V. Van Speybroeck, and D. Van Neck, “Communication: DMRG-SCF study of the singlet, triplet, and quintet states of oxo-Mn(Salen)”, *The Journal of Chemical Physics* **140**, 241103 (2014).
- [145] Y. Kurashige, G. K.-L. Chan, and T. Yanai, “Entangled quantum electronic wavefunctions of the Mn_4CaO_5 cluster in photosystem II”, *Nature Chemistry* **5**, 660–666 (2013).
- [146] K. Boguslawski, P. Tecmer, Ö. Legeza, and M. Reiher, “Entanglement measures for single- and multireference correlation effects”, *The Journal of Physical Chemistry Letters* **3**, 3129–3135 (2012).

- [147] K. Boguslawski, K. H. Marti, Ö. Legeza, and M. Reiher, “Accurate ab initio spin densities”, *Journal of Chemical Theory and Computation* **8**, 1970–1982 (2012).
- [148] S. Sharma and G. K.-L. Chan, “Spin-adapted density matrix renormalization group algorithms for quantum chemistry”, *The Journal of Chemical Physics* **136**, 124121 (2012).
- [149] G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher, “Quantum-information analysis of electronic states of different molecular structures”, *Physical Review A* **83**, 012508 (2011).
- [150] T. Yanai, Y. Kurashige, E. Neuscamman, and G. K.-L. Chan, “Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory”, *The Journal of Chemical Physics* **132**, 024105 (2010).
- [151] Y. Kurashige and T. Yanai, “High-performance ab initio density matrix renormalization group method: applicability to large-scale multireference problems for metal compounds”, *The Journal of Chemical Physics* **130**, 234114 (2009).
- [152] K. H. Marti, I. M. Ondík, G. Moritz, and M. Reiher, “Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters”, *The Journal of Chemical Physics* **128**, 014104 (2008).
- [153] G. Moritz and M. Reiher, “Construction of environment states in quantum-chemical density-matrix renormalization group calculations”, *The Journal of Chemical Physics* **124**, 034103 (2006).
- [154] J. Brabec, J. Brandejs, K. Kowalski, S. Xantheas, Ö. Legeza, and L. Veis, “Massively parallel quantum chemical density matrix renormalization group method”, arXiv e-prints, 2001.04890v1 (2020).
- [155] H. R. Larsson, C. A. Jiménez-Hoyos, and G. K.-L. Chan, “Minimal matrix product states and generalizations of mean-field and geminal wave functions”, *Journal of Chemical Theory and Computation* **16**, 5057–5066 (2020).
- [156] K. Gunst, *T3NS: an implementation of the three-legged tree tensor network algorithm*, <https://github.com/klgunst/T3NS>, 2020.
- [157] E. R. Davidson, “The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices”, *Journal of Computational Physics* **17**, 87–94 (1975).

- [158] J. LeBlanc et al., “Solutions of the two-dimensional Hubbard model: benchmarks and results from a wide range of numerical algorithms”, *Physical Review X* **5**, 041041 (2015).
- [159] H.-C. Jiang, H. Yao, and L. Balents, “Spin liquid ground state of the spin- $\frac{1}{2}$ square $J_1 - J_2$ Heisenberg model”, *Physical Review B* **86**, 024424 (2012).
- [160] S. Yan, D. A. Huse, and S. R. White, “Spin-liquid ground state of the $S = 1/2$ Kagome Heisenberg antiferromagnet”, *Science* **332**, 1173–1176 (2011).
- [161] S. R. White, “Spin gaps in a frustrated heisenberg model for CaV_4O_9 ”, *Physical Review Letters* **77**, 3633–3636 (1996).
- [162] F. Verstraete and J. I. Cirac, “Renormalization algorithms for quantum-many body systems in two and higher dimensions”, arXiv e-prints, cond-mat/0407066 (2004).
- [163] M. P. Zaletel and F. Pollmann, “Isometric tensor network states in two dimensions”, *Physical Review Letters* **124**, 037201 (2020).
- [164] G. Evenbly, “Gauge fixing, canonical forms, and optimal truncations in tensor networks with closed loops”, *Physical Review B* **98**, 085155 (2018).
- [165] R. Haghshenas, M. J. O’Rourke, and G. K.-L. Chan, “Conversion of projected entangled pair states into a canonical form”, *Physical Review B* **100**, 054404 (2019).
- [166] G. Vidal, “Entanglement renormalization”, *Physical Review Letters* **99**, 220405 (2007).
- [167] G. Evenbly and G. Vidal, “Tensor network states and geometry”, *Journal of Statistical Physics* **145**, 891–918 (2011).
- [168] V. Murg, F. Verstraete, Ö. Legeza, and R. M. Noack, “Simulating strongly correlated quantum systems with tree tensor networks”, *Physical Review B* **82**, 205105 (2010).
- [169] N. Nakatani and G. K.-L. Chan, “Efficient tree tensor network states (TTNS) for quantum chemistry: generalizations of the density matrix renormalization group algorithm”, *The Journal of Chemical Physics* **138**, 134113 (2013).
- [170] V. Murg, F. Verstraete, R. Schneider, P. R. Nagy, and Ö. Legeza, “Tree tensor network state with variable tensor order: an efficient multireference method for strongly correlated systems”, *Journal of Chemical Theory and Computation* **11**, 1027–1036 (2015).

- [171] K. H. Marti, B. Bauer, M. Reiher, M. Troyer, and F. Verstraete, “Complete-graph tensor network states: a new fermionic wave function ansatz for molecules”, *New Journal of Physics* **12**, 103008 (2010).
- [172] A. Kovyrshin and M. Reiher, “Self-adaptive tensor network states with multi-site correlators”, *The Journal of Chemical Physics* **147**, 214111 (2017).
- [173] F. Verstraete and J. I. Cirac, “Continuous matrix product states for quantum fields”, *Physical Review Letters* **104**, 190405 (2010).
- [174] B. Tuybens, J. De Nardis, J. Haegeman, and F. Verstraete, “Variational optimization of continuous matrix product states”, arXiv e-prints, 2006.01801 (2020).
- [175] J. Haegeman, T. J. Osborne, H. Verschelde, and F. Verstraete, “Entanglement renormalization for quantum fields in real space”, *Physical Review Letters* **110**, 100402 (2013).
- [176] S. Wouters, W. Poelmans, P. W. Ayers, and D. Van Neck, “CheMPS2: a free open-source spin-adapted implementation of the density matrix renormalization group for ab initio quantum chemistry”, *Computer Physics Communications* **185**, 1501–1514 (2014).
- [177] Y.-Y. Shi, L.-M. Duan, and G. Vidal, “Classical simulation of quantum many-body systems with a tree tensor network”, *Physical Review A* **74**, 022320 (2006).
- [178] A. J. Ferris, “Area law and real-space renormalization”, *Physical Review B* **87**, 125139 (2013).
- [179] M. Thoss, H. Wang, and W. H. Miller, “Self-consistent hybrid approach for complex systems: application to the spin-boson model with Debye spectral density”, *The Journal of Chemical Physics* **115**, 2991–3005 (2001).
- [180] U. Manthe, “A multilayer multiconfigurational time-dependent Hartree approach for quantum dynamics on general potential energy surfaces”, *The Journal of Chemical Physics* **128**, 164116 (2008).
- [181] H. Wang and M. Thoss, “Numerically exact quantum dynamics for indistinguishable particles: the multilayer multiconfiguration time-dependent Hartree theory in second quantization representation”, *The Journal of Chemical Physics* **131**, 024114 (2009).
- [182] Ö. Legeza, J. Röder, and B. A. Hess, “Controlling the accuracy of the density-matrix renormalization-group method: the dynamical block state selection approach”, *Physical Review B* **67**, 125114 (2003).

- [183] L. Tagliacozzo, G. Evenbly, and G. Vidal, “Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law”, *Physical Review B* **80**, 235127 (2009).
- [184] S. Cheng, L. Wang, T. Xiang, and P. Zhang, “Tree tensor networks for generative modeling”, *Physical Review B* **99**, 155131 (2019).
- [185] Ö. Legeza and J. Sólyom, “Optimizing the density-matrix renormalization group method using quantum information entropy”, *Physical Review B* **68**, 195116 (2003).
- [186] J. Rissler, R. M. Noack, and S. R. White, “Measuring orbital interaction using quantum information theory”, *Chemical Physics* **323**, 519–531 (2006).
- [187] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, “Orbital optimization in the density matrix renormalization group, with applications to polyenes and β -carotene”, *The Journal of Chemical Physics* **128**, 144117 (2008).
- [188] D. Zgid and M. Nooijen, “The density matrix renormalization group self-consistent field method: orbital optimization with the density matrix renormalization group method in the active space”, *The Journal of Chemical Physics* **128**, 144116 (2008).
- [189] C. Krumnow, L. Veis, Ö. Legeza, and J. Eisert, “Fermionic orbital optimization in tensor network states”, *Physical Review Letters* **117**, 210402 (2016).
- [190] P. Jordan and E. P. Wigner, “Über das Paulische Äquivalenzverbot”, in *The Collected Works of Eugene Paul Wigner: Part A: The Scientific Papers*, edited by A. S. Wightman (Springer Berlin Heidelberg, Berlin, Heidelberg, 1993), pp. 109–129.
- [191] N. Bultinck, D. J. Williamson, J. Haegeman, and F. Verstraete, “Fermionic matrix product states and one-dimensional topological phases”, *Physical Review B* **95**, 075108 (2017).
- [192] S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza, “Tensor product methods and entanglement optimization for *ab initio* quantum chemistry”, *International Journal of Quantum Chemistry* **115**, 1342–1391 (2015).
- [193] S. Wouters and D. Van Neck, “The density matrix renormalization group for *ab initio* quantum chemistry”, *The European Physical Journal D* **68**, 272 (2014).

- [194] C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, “Theoretical models on the Cu_2O_2 torture track: mechanistic implications for oxytyrosinase and small-molecule analogues”, *The Journal of Physical Chemistry A* **110**, 1991–2004 (2006).
- [195] P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, “The restricted active space followed by second-order perturbation theory method: theory and application to the study of CuO_2 and Cu_2O_2 systems”, *The Journal of Chemical Physics* **128**, 204109 (2008).
- [196] K. H. Marti and M. Reiher, “The density matrix renormalization group algorithm in quantum chemistry”, *Zeitschrift für Physikalische Chemie* **224**, 583–599 (2010).
- [197] C. J. Stein and M. Reiher, “Automated selection of active orbital spaces”, *Journal of Chemical Theory and Computation* **12**, 1760–1771 (2016).
- [198] C. W. Bauschlicher and S. R. Langhoff, “Full configuration-interaction study of the ionic-neutral curve crossing in LiF ”, *The Journal of Chemical Physics* **89**, 4246–4254 (1988).
- [199] G. K.-L. Chan, M. Kállay, and J. Gauss, “State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve”, *The Journal of Chemical Physics* **121**, 6110–6116 (2004).
- [200] H. Larsen, J. Olsen, P. Jørgensen, and O. Christiansen, “Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of N_2 ”, *The Journal of Chemical Physics* **113**, 6677–6686 (2000).
- [201] M. Dolg, U. Wedig, H. Stoll, and H. Preuss, “Energy-adjusted ab initio pseudopotentials for the first row transition elements”, *The Journal of Chemical Physics* **86**, 866–872 (1987).
- [202] K. Pierloot, B. Dumez, P.-O. Widmark, and B. O. Roos, “Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions”, *Theoretica Chimica Acta* **90**, 87–114 (1995).
- [203] A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, “Quantum chemistry using the density matrix renormalization group”, *The Journal of Chemical Physics* **115**, 6815–6821 (2001).

- [204] G. K.-L. Chan and M. Head-Gordon, “Highly correlated calculations with a polynomial cost algorithm: a study of the density matrix renormalization group”, *The Journal of Chemical Physics* **116**, 4462–4476 (2002).
- [205] G. Sierra and T. Nishino, “The density matrix renormalization group method applied to interaction round a face Hamiltonians”, *Nuclear Physics B* **495**, 505–532 (1997).
- [206] I. P. McCulloch and M. Gulácsi, “Density matrix renormalisation group method and symmetries of the Hamiltonian”, *Australian Journal of Physics* **53**, 597 (2000).
- [207] I. P. McCulloch and M. Gulácsi, “Total spin in the density matrix renormalization group algorithm”, *Philosophical Magazine Letters* **81**, 447–453 (2001).
- [208] I. P. McCulloch and M. Gulácsi, “The non-abelian density matrix renormalization group algorithm”, *Europhysics Letters* **57**, 852–858 (2002).
- [209] E. P. Wigner, *Group theory: and its application to the quantum mechanics of atomic spectra*, Vol. 5, Pure and Applied Physics (Academic Press, New York, NY, 1959).
- [210] I. P. McCulloch, “From density-matrix renormalization group to matrix product states”, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, P10014–P10014 (2007).
- [211] D. Zgid and M. Nooijen, “On the spin and symmetry adaptation of the density matrix renormalization group method”, *The Journal of Chemical Physics* **128**, 014107 (2008).
- [212] S. Keller and M. Reiher, “Spin-adapted matrix product states and operators”, *The Journal of Chemical Physics* **144**, 134101 (2016).
- [213] A. I. Tóth, C. P. Moca, Ö. Legeza, and G. Zaránd, “Density matrix numerical renormalization group for non-abelian symmetries”, *Physical Review B* **78**, 245109 (2008).
- [214] S. Singh and G. Vidal, “Tensor network states and algorithms in the presence of a global SU(2) symmetry”, *Physical Review B* **86**, 195114 (2012).
- [215] C. Hubig, “Abelian and non-abelian symmetries in infinite projected entangled pair states”, *SciPost Physics* **5**, 47 (2018).
- [216] P. Schmoll, S. Singh, M. Rizzi, and R. Orús, “A programming guide for tensor networks with global su(2) symmetry”, *Annals of Physics* **419**, 168232 (2020).

- [217] S. Sharma, “A general non-abelian density matrix renormalization group algorithm with application to the C_2 dimer”, *The Journal of Chemical Physics* **142**, 024107 (2015).
- [218] A. Weichselbaum, “Non-abelian symmetries in tensor networks: a quantum symmetry space approach”, *Annals of Physics* **327**, 2972–3047 (2012).
- [219] S. Singh, R. N. C. Pfeifer, and G. Vidal, “Tensor network decompositions in the presence of a global symmetry”, *Physical Review A* **82**, 050301 (2010).
- [220] S. Singh, H.-Q. Zhou, and G. Vidal, “Simulation of one-dimensional quantum systems with a global $SU(2)$ symmetry”, *New Journal of Physics* **12**, 033029 (2010).
- [221] K. Gunst, D. Van Neck, P. A. Limacher, and S. De Baerdemacker, “The seniority quantum number in tensor network states”, arXiv e-prints, 2008.00733 (2020).
- [222] G. Racah, “Theory of complex spectra. III”, *Physical Review* **63**, 367–382 (1943).
- [223] P. Ring and P. Schuck, *The nuclear many-body problem* (Springer, 1980).
- [224] F. Weinhold and E. B. Wilson, “Reduced density matrices of atoms and molecules. I. The 2 matrix of double-occupancy, configuration-interaction wavefunctions for singlet states”, *The Journal of Chemical Physics* **46**, 2752–2758 (1967).
- [225] F. Weinhold and E. B. Wilson, “Reduced density matrices of atoms and molecules. II. On the N-representability problem”, *The Journal of Chemical Physics* **47**, 2298–2311 (1967).
- [226] L. Bytautas, T. M. Henderson, C. A. Jiménez-Hoyos, J. K. Ellis, and G. E. Scuseria, “Seniority and orbital symmetry as tools for establishing a full configuration interaction hierarchy”, *The Journal of Chemical Physics* **135**, 044119 (2011).
- [227] W. Kutzelnigg, “Separation of strong (bond-breaking) from weak (dynamical) correlation”, *Chemical Physics* **401**, 119–124 (2012).
- [228] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, “A new mean-field method suitable for strongly correlated electrons: computationally facile antisymmetric products of nonorthogonal geminals”, *Journal of Chemical Theory and Computation* **9**, 1394–1401 (2013).

- [229] D. R. Alcoba, A. Torre, L. Lain, G. E. Massaccesi, and O. B. Oña, “Configuration interaction wave functions: a seniority number approach”, *The Journal of Chemical Physics* **140**, 234103 (2014).
- [230] P. A. Limacher, “A new wavefunction hierarchy for interacting geminals”, *The Journal of Chemical Physics* **145**, 194102 (2016).
- [231] K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, “Efficient description of strongly correlated electrons with mean-field cost”, *Physical Review B* **89**, 201106 (2014).
- [232] K. Boguslawski, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, “Projected seniority-two orbital optimization of the antisymmetric product of one-reference orbital geminal”, *The Journal of Chemical Physics* **140**, 214114 (2014).
- [233] K. Boguslawski, P. Tecmer, P. Bultinck, S. De Baerdemacker, D. Van Neck, and P. W. Ayers, “Nonvariational orbital optimization techniques for the AP1roG wave function”, *Journal of Chemical Theory and Computation* **10**, 4873–4882 (2014).
- [234] T. Stein, T. M. Henderson, and G. E. Scuseria, “Seniority zero pair coupled cluster doubles theory”, *The Journal of Chemical Physics* **140**, 214113 (2014).
- [235] T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria, “Seniority-based coupled cluster theory”, *The Journal of Chemical Physics* **141**, 244104 (2014).
- [236] J. J. Shepherd, T. M. Henderson, and G. E. Scuseria, “Using full configuration interaction quantum monte carlo in a seniority zero space to investigate the correlation energy equivalence of pair coupled cluster doubles and doubly occupied configuration interaction”, *The Journal of Chemical Physics* **144**, 094112 (2016).
- [237] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, “Simple and inexpensive perturbative correction schemes for antisymmetric products of nonorthogonal geminals”, *Physical Chemistry Chemical Physics* **16**, 5061 (2014).
- [238] J. M. Wahlen-Strothman, T. M. Henderson, and G. E. Scuseria, “Influence of broken-pair excitations on the exact pair wavefunction”, *Molecular Physics* **116**, 186–193 (2017).
- [239] L. Bytautas, G. E. Scuseria, and K. Ruedenberg, “Seniority number description of potential energy surfaces: symmetric dissociation of water, N₂, C₂, and Be₂”, *The Journal of Chemical Physics* **143**, 094105 (2015).

- [240] L. Bytautas and J. Dukelsky, “Seniority based energy renormalization group (Ω -ERG) approach in quantum chemistry: initial formulation and application to potential energy surfaces”, *Computational and Theoretical Chemistry* **1141**, 74–88 (2018).
- [241] Q. Sun et al., “Recent developments in the PySCF program package”, *The Journal of Chemical Physics* **153**, 024109 (2020).
- [242] Q. Sun et al., “PySCF: the python-based simulations of chemistry framework”, *WIREs Computational Molecular Science* **8**, e1340 (2017).
- [243] Q. Sun, “Libcint: an efficient general integral library for Gaussian basis functions”, *Journal of Computational Chemistry* **36**, 1664–1671 (2015).
- [244] P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, “The influence of orbital rotation on the energy of closed-shell wavefunctions”, *Molecular Physics* **112**, 853–862 (2014).
- [245] R. A. Aziz and M. Slaman, “The Ne-Ne interatomic potential revisited”, *Chemical Physics* **130**, 187–194 (1989).
- [246] T. Helgaker, P. Jørgensen, and J. Olsen, “Gaussian basis sets”, in *Molecular electronic-structure theory* (John Wiley & Sons, Ltd, 2014) Chap. 8, pp. 287–335.
- [247] S. Boys and F. Bernardi, “The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors”, *Molecular Physics* **19**, 553–566 (1970).
- [248] G. Moritz, B. A. Hess, and M. Reiher, “Convergence behavior of the density-matrix renormalization group algorithm for optimized orbital orderings”, *The Journal of Chemical Physics* **122**, 024107 (2005).
- [249] G. K.-L. Chan and S. Sharma, “The density matrix renormalization group in quantum chemistry”, *Annual Review of Physical Chemistry* **62**, 465–481 (2011).
- [250] A. O. Mitrushchenkov, G. Fano, R. Linguerri, and P. Palmieri, “On the importance of orbital localization in QC-DMRG calculations”, *International Journal of Quantum Chemistry* **112**, 1606–1619 (2012).
- [251] E. Fertitta, B. Paulus, G. Barcza, and Ö. Legeza, “Investigation of metal–insulator-like transition through the ab initio density matrix renormalization group approach”, *Physical Review B* **90**, 245129 (2014).

- [252] J. Hachmann, W. Cardoen, and G. K.-L. Chan, “Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group”, *The Journal of Chemical Physics* **125**, 144101 (2006).
- [253] S. Wouters, P. A. Limacher, D. Van Neck, and P. W. Ayers, “Longitudinal static optical properties of hydrogen chains: finite field extrapolations of matrix product state calculations”, *The Journal of Chemical Physics* **136**, 134110 (2012).
- [254] W. Mizukami, Y. Kurashige, and T. Yanai, “More π electrons make a difference: emergence of many radicals on graphene nanoribbons studied by ab initio DMRG theory”, *Journal of Chemical Theory and Computation* **9**, 401–407 (2013).
- [255] J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, “The radical character of the acenes: a density matrix renormalization group study”, *The Journal of Chemical Physics* **127**, 134309 (2007).
- [256] J. J. Dorando, J. Hachmann, and G. K.-L. Chan, “Targeted excited state algorithms”, *The Journal of Chemical Physics* **127**, 084109 (2007).
- [257] N. Nakatani and S. Guo, “Density matrix renormalization group (DMRG) method as a common tool for large active-space CASSCF/CASPT2 calculations”, *The Journal of Chemical Physics* **146**, 094102 (2017).
- [258] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, “The ab-initio density matrix renormalization group in practice”, *The Journal of Chemical Physics* **142**, 034102 (2015).
- [259] S. Sharma, K. Sivalingam, F. Neese, and G. K.-L. Chan, “Low-energy spectrum of iron–sulfur clusters directly from many-particle quantum mechanics”, *Nature Chemistry* **6**, 927–933 (2014).
- [260] Z. Li, S. Guo, Q. Sun, and G. K.-L. Chan, “Electronic landscape of the P-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations”, *Nature Chemistry* **11**, 1026–1033 (2019).
- [261] S. F. Keller and M. Reiher, “Determining factors for the accuracy of DMRG in chemistry”, *CHIMIA International Journal for Chemistry* **68**, 200–203 (2014).
- [262] E. M. Stoudenmire and S. R. White, “Sliced basis density matrix renormalization group for electronic structure”, *Physical Review Letters* **119**, 046401 (2017).

- [263] M. Motta et al. (Simons Collaboration on the Many-Electron Problem), “Towards the solution of the many-electron problem in real materials: equation of state of the hydrogen chain with state-of-the-art many-body methods”, *Physical Review X* **7**, 031059 (2017).
- [264] S. R. White, “Hybrid grid/basis set discretizations of the Schrödinger equation”, *The Journal of Chemical Physics* **147**, 244102 (2017).
- [265] S. R. White and E. M. Stoudenmire, “Multisliced gausslet basis sets for electronic structure”, *Physical Review B* **99**, 081110 (2019).
- [266] C. Krumnow, L. Veis, J. Eisert, and Ö. Legeza, “Dimension reduction with mode transformations: simulating two-dimensional fermionic condensed matter systems”, arXiv e-prints, 1906.00205 (2019).
- [267] C. Krumnow, J. Eisert, and Ö. Legeza, “Towards overcoming the entanglement barrier when simulating long-time evolution”, arXiv e-prints, 1904.11999 (2019).
- [268] J. E. Atkins, E. G. Boman, and B. Hendrickson, “A spectral algorithm for seriation and the consecutive ones problem”, *SIAM Journal on Computing* **28**, 297–310 (1998).
- [269] E. Cuthill and J. McKee, “Proceedings of the 24th national conference of the association for computing machinery”, (1969).
- [270] W.-H. Liu and A. H. Sherman, “Comparative analysis of the Cuthill–McKee and the reverse Cuthill–McKee ordering algorithms for sparse matrices”, *SIAM Journal on Numerical Analysis* **13**, 198–213 (1976).
- [271] Ö. Legeza, J. Röder, and B. A. Hess, “QC-DMRG study of the ionic-neutral curve crossing of LiF”, *Molecular Physics* **101**, 2019–2028 (2003).
- [272] A. O. Mitrushenkov, R. Linguerri, P. Palmieri, and G. Fano, “Quantum chemistry using the density matrix renormalization group II”, *The Journal of Chemical Physics* **119**, 4148–4158 (2003).
- [273] Ö. Legeza and J. Sólyom, “Quantum data compression, quantum information generation, and the density-matrix renormalization-group method”, *Physical Review B* **70**, 205118 (2004).
- [274] Ö. Legeza, L. Veis, A. Poves, and J. Dukelsky, “Advanced density matrix renormalization group method for nuclear structure calculations”, *Physical Review C* **92**, 051303 (2015).

- [275] Z. Luo, Y. Ma, C. Liu, and H. Ma, “Efficient reconstruction of CAS-CI-type wave functions for a DMRG state using quantum information theory and a genetic algorithm”, *Journal of Chemical Theory and Computation* **13**, 4699–4710 (2017).
- [276] H. Araki and E. H. Lieb, “Entropy inequalities”, *Communications in Mathematical Physics*, 160–170 (1970).
- [277] K. M. R. Audenaert, “Subadditivity of q -entropies for $q > 1$ ”, *Journal of Mathematical Physics* **48**, 083507 (2007).
- [278] G. A. Raggio, “Properties of q -entropies”, *Journal of Mathematical Physics* **36**, 4785–4791 (1995).
- [279] H. R. Larsson, “Computing vibrational eigenstates with tree tensor network states (TTNS)”, *The Journal of Chemical Physics* **151**, 204102 (2019).
- [280] C. Eckart and G. Young, “The approximation of one matrix by another of lower rank”, *Psychometrika* **1**, 211–218 (1936).
- [281] G. W. Stewart, “On the early history of the singular value decomposition”, *SIAM Review* **35**, 551–566 (1993).
- [282] L. Mirsky, “Symmetric gauge functions and unitarily invariant norms”, *The Quarterly Journal of Mathematics* **11**, 50–59 (1960).
- [283] H. A. Jahn and J. Hope, “Symmetry properties of the Wigner $9j$ symbol”, *Physical Review* **93**, 318–321 (1954).