

RICE UNIVERSITY

**Lie algebraic similarity transformations: improving
wavefunctions for weak and strong correlations**

by

Jacob M. Wahlen-Strothman

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

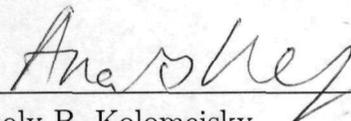
APPROVED, THESIS COMMITTEE:



Gustavo E. Scuseria, Chair
Robert A. Welch Professor of Chemistry,
Physics and Astronomy and Materials
Science and NanoEngineering



Andriy Nevidomskyy
Assistant Professor of Physics and
Astronomy



Anatoly B. Kolomeisky
Professor of Chemistry and Chemical and
Biomolecular Engineering

Houston, Texas

May, 2017

ABSTRACT

Lie algebraic similarity transformations: improving wavefunctions for weak and strong correlations

by

Jacob M. Wahlen-Strothman

We present a class of correlated wavefunctions generated by exponentials of two-body on-site Hermitian operators that can be evaluated with polynomial computational cost via a Hamiltonian similarity transformation. Wavefunctions of this form have been studied with variational Monte Carlo methods, but we present a formalism to perform non-stochastic calculations. The Hausdorff series generated by these Jastrow factors can be summed exactly without truncation resulting in a set of equations with polynomial computational cost. The correlators include the density-density, collinear spin-spin, spin-density cross terms, and on-site double occupancy operators. The resulting non-Hermitian many-body Hamiltonian can be solved in a biorthogonal mean-field approach with only a small set of correlation terms required for accurate calculations in systems with local interactions. Although the energy of the model is unbound, projective equations in the spirit of coupled cluster theory lead to well-defined solutions. The theory is tested on the one and two-dimensional repulsive Hubbard model where it yields accurate results for large systems with low computational cost. Symmetry projection methods are included to further improve the reference wavefunction and results under strong correlation without sacrificing good quantum numbers resulting in very accurate energies for small systems and

producing a better ground state for the calculation of other properties.

Acknowledgments

I would first like to thank my advisor Dr. Gustavo E. Scuseria for guidance through this research. There are also several current and former group members who have provided invaluable support and shared their experience in my time at Rice. Dr. Carlos Jiménez-Hoyos, Dr Thomas Henderson, and Dr. Ireneusz W. Bulik have been incredibly supportive providing insights, valuable discussion, and extensive assistance in new topics. I would like to thank Jin-Mo Zhao for early exploratory work that inspired the early direction of this research project. I would also like to thank Dr. Matthias Degroote, Dr. Matthew Hermes, and John A. Gomez for a number of insightful discussions over the years.

Contents

Abstract	ii
Acknowledgments	iv
List of Illustrations	vii
List of Tables	ix
Preface	xi
1 Introduction	1
1.1 Jastrow correlated wavefunctions	2
1.2 Weak and Strong Correlation	3
2 Background Material	6
2.1 Notation	6
2.2 Hamiltonians	7
2.2.1 The Hubbard Model	7
2.2.2 General Electronic Hamiltonians	8
2.3 Hartree-Fock	9
2.4 Symmetry Breaking and Restoration	16
2.4.1 Symmetry Broken Hartree-Fock	16
2.4.2 Projected Hartree-Fock	18
3 The Lie Algebraic Similarity Transformation (LAST)	20
3.1 Jastrow-type Correlation Factors	21
3.2 Evaluating the Similarity Transformation	23
3.3 The Working Equations	25

3.4	LAST on the Hubbard Model	27
4	Improving the Reference Wavefunction	33
4.1	The Calculation Scheme	34
4.1.1	Orbital Optimization	35
4.1.2	Using a PHF Reference	37
4.2	Results and Discussions	39
5	LAST for General Hamiltonians	45
5.1	The Gutzwiller Similarity Transformed Hamiltonian	45
5.2	Orbital, and Basis Optimization Equations	49
6	Conclusions	53
A	Evaluation of PHF Expectation Values	55
B	Additional Expressions and Results for LAST	57
B.1	The LAST transformation	57
B.2	LAST Expectation Values	58
B.3	GST Expectation Values	60
B.4	Additional Results	62
	Bibliography	68

Illustrations

2.1	Mean-field energies of a 10-site Hubbard ring as a function of the interaction strength.	17
3.1	Correlation energy of eight-hole doped Hubbard chains for $U = 2$ with open and closed boundaries on an RHF reference. Density matrix renormalization group (DMRG) is used to find exact energies for open systems [34,35].	28
3.2	The spin-spin correlation function in Fourier space calculated using Equation 3.25 (S_J) for a 30-site Hubbard ring at half filling and $U = 3$ compared to the RHF reference and DMRG. [34,35].	29
3.3	The cumulative fraction of correlation energy captured by limiting the range R of the correlators compared to the full set (E_F) for a 70-site ring (a), and 8×8 2D lattice (b), at half filling with an RHF reference. $R = 0$ includes the Gutzwiller factor.	31
4.1	Error in the energy per site for 4×4 square Hubbard lattices with 16 particles (left) and 14 particles (right) compared to exact diagonalization from Ref. [38].	39
4.2	Time required for the UGST orbital optimization and SUGST amplitude optimization with integration grid size equal to $\sqrt{N_{sites}}$ on half-filled Hubbard square lattices with $U = 4$	40

4.3	Energy errors and finite-size effects per site for 10×10 square Hubbard lattices with $\langle n \rangle = 1$ (left) and $\langle n \rangle = 0.8$ (right) compared to UCCSD and averages of high quality thermodynamic limit calculations (E_{TDL}) from Refs. [39,40] as accurate results for finite systems are limited.	40
4.4	Energies per site for the half-filled Hubbard model approaching the thermodynamic limit for N -site square lattices at $U = 4$, with an average of high quality results for an infinite system (TDL) from Ref. [39].	41
4.5	Energy errors per site for 30×30 square Hubbard lattices with $\langle n \rangle = 1$ (left) and $\langle n \rangle = 0.8$ (right) compared to extrapolated variational Monte Carlo (VMC) from Ref. [41] as well as extrapolated UCCSD and averages of high quality thermodynamic limit calculations (E_{TDL}) from Ref. [39].	42
4.6	Double occupancy errors per site for 30×30 square Hubbard lattices with $\langle n \rangle = 1$ compared to UCCSD and average of high quality thermodynamic limit calculations (D_{TDL}) from Ref. [39].	42
4.7	Spin-spin correlation function with alternating sign for a 30-site Hubbard ring with $U = 4$ compared to the exact density matrix renormalization group theory (DMRG) result. [34,35]	44
5.1	H_4 square with dimensions $1 \times \epsilon$ Angstroms dissociating into two H_2 dimers using the GST in the 3-21G basis including RHF as the reference (RGST), orbital optimization (oo-RGST), and orbital and correlator basis optimization (oo-bo-RGST).	52

Tables

3.1	Energy per site and portion of the recovered correlation energy (E_c) for 2D, periodic lattices with occupancy N_o , an RHF reference and released-constraint Monte Carlo (E_{MC}) [36,37] as the best estimate for the exact result.	29
3.2	Energies and overlaps for the exact $ 0\rangle$, RHF $ \Phi\rangle$, and correlated wavefunctions $ J\rangle$ for a 14-site ring, where $ J\rangle = e^J \Phi\rangle/ \langle\Phi e^{2J} \Phi\rangle ^{\frac{1}{2}}$.	30
3.3	Energies for 4×4 Hubbard lattices with RHF (E_{RJ}) and UHF (E_{UJ}) references including spin-density correlators ($S_i^z N_j$), compared to exact energies (E_{ED}) [36,38].	31
4.1	Gutzwiller correlation factors (g) and antiferromagnetic order parameters (M) from SUGST on a 30×30 , half-filled lattice and variational Monte Carlo (VMC) extrapolated to the thermodynamic limit.	43
B.1	Correlated energies for 4×4 Hubbard lattices on RHF (E_{RJ}) and UHF (E_{UJ}) references and the portion of the correlation energy recovered (E_c), exact energies (E_{ED}) taken from [36]. E_{UJ} includes spin-density correlators ($S_i^z N_j$).	63
B.2	Correlation energies per site with respect to UHF for 4×4 Hubbard lattices with spin quantum numbers $s = m = 0$ and N_o electrons. Exact energies (ED) taken from [38].	64

B.3	Correlation energies per site with respect to UHF for 10×10 lattices with average occupancy n	65
B.4	SUGST correlation energies per site for 30×30 square lattices with average occupancy n	66
B.5	Average SUGST double occupancies per site for 30×30 square lattices at half-filling.	67

Preface

The work outlined in this manuscript represents the results of a project to develop a new method to perform calculations in many-body quantum mechanical systems. This work was done from 2012 to 2016 under the supervision of Dr. Gustavo E. Scuseria. Some of the results presented in this work have been published in the following articles:

- *Lie algebraic similarity transformed Hamiltonians for lattice model systems*, J. M. Wahlen-Strothman, C. A. Jiménez-Hoyos, T. M. Henderson and G. E. Scuseria, *Phys. Rev. B*, **91**, 041114(R) (2015).
- *Biorthogonal projected energies of a Gutzwiller similarity transformed Hamiltonian*, J. M. Wahlen-Strothman and G. E. Scuseria, *J. Phys. Condens. Matt.* **28**, 485502 (2016).

In addition there are some further results that will be discussed later on. This work was inspired by previous results found in:

- *Nonstochastic algorithms for Jastrow-Slater and correlator product state wave functions*, E. Neuscamman, H. Changlani, J. Kinder, and G. K.-L. Chan, *Phys. Rev. B*, **84**, 205132 (2011).
- *Transcorrelated method: Another possible way towards electronic structure calculation of solids*, S. Tsuneyuki, *Progress of Theoretical Physics Supplement*, **176**, 134 (2008).

There has been additional, similar work done with other methods of evaluation in a number of works including:

- *Communication: A jastrow factor coupled cluster theory for weak and strong electron correlation*, E. Neuscamman, *The Journal of Chemical Physics*, **139**, 181101 (2013).
- *Variational monte-carlo studies of hubbard model. ii*, H. Yokoyama and H. Shiba, *Journal of the Physical Society of Japan*, **56**, 3582 (1987).

Chapter 1

Introduction

Useful wavefunction approximations are a valuable tool for many-body quantum mechanics. There is great interest in producing accurate calculations to both guide and supplement experimentation on new chemicals and materials. However, attempts to produce these theoretical predictions must always overcome the fundamental limitation of many-body calculations, the size of the Hilbert space. As the number of parameters to describe the true solution increases exponentially with system size, careful approximations must be implemented both to decrease the degrees of freedom and the computational complexity to reasonable sizes.

In this manuscript, we will explore a new method for evaluating a form of correlated wavefunctions that effectively treats a wide range of correlation strengths. These wavefunctions have long been popular in the Monte Carlo community. Here we evaluate them via a similarity transformation method resulting in a resumable series with polynomial computational cost and no need for stochastic sampling. The low cost of these calculations allows us to benchmark very large systems and the form of the equations permits the consideration of more advanced wavefunctions beyond simple mean-field.

1.1 Jastrow correlated wavefunctions

While many of the methods that improve on the mean-field calculations can be described as building additional correlations into a mean-field reference, there is a continuous search for new methods of introducing relevant correlations to a system. These efforts are made difficult as the developed methods must produce accurate results while incorporating a reasonable number of parameters and maintaining a low computational cost. In this work we examine the usefulness of a two-body Jastrow factor to build correlations into the solution. The wavefunctions take the form,

$$|\Psi\rangle = e^J|\Phi\rangle, \quad (1.1)$$

where $|\Phi\rangle$ is a mean-field wavefunction that is easy to perform calculations with, and J is a set of Jastrow-type correlation factors that build more advanced structure into the solution. The components of this operator are directly constructed out of two-point correlation factors such as spin-spin ($S_i^z S_j^z$), density-density ($N_i N_j$) and double occupancy ($n_{i\uparrow} n_{i\downarrow}$) operators. These have the advantage that they contain few parameters when compared to other common methods, such as coupled cluster, but they hold much of the relevant physics as they are constructed to directly modify important correlations within the system.

The operators considered here are not new, and they have been implemented as correlation factors in the past with methods such as variational Monte Carlo. In this work we implement them with a new approach requiring no stochastic sampling. The resulting methods provide low-cost approximations that capture relevant physics of both weak and strongly correlated systems.

1.2 Weak and Strong Correlation

Most correlated methods for electronic structure can be broken up into one of three main categories. These groups are classified by the types of problems each method is effective for and where they break down.

The first and most common category includes methods that are effective for weak, or dynamic, correlation. Weakly correlated systems are those where the positions of electrons have little dependence on the positions of all the other electrons in the system. This could be the case when the particles are non-interacting, or in systems such as molecules where the electron density can be distributed uniformly. Weakly correlated systems are generally well approximated by a set of independent particles placed in a set of optimized, single particle states (sometimes called the molecular or canonical orbitals), and as such mean field approximations are reasonable and a good starting point for correlated methods. Methods effective for these systems, such as coupled cluster [1] and density functional theory, are most common as many molecules are dominated by weak correlations at equilibrium geometries.

The second category of methods includes those effective for strongly correlated systems. Strong, or static, correlation is generally associated with strong two-body interactions, degeneracies, and long range collective behavior. The behavior of the electrons in these systems is highly dependent on the positions of the other electrons. For example in a state with antiferromagnetic character if we measure the spin polarization of one site, we know with some certainty that the neighboring sites will have the opposite polarization. In these systems the single particle picture breaks down as any single independent particle will have low occupation in the exact solution. Again with the antiferromagnetic example, there is no single mean-field solution describes the Néel structure while maintaining good spin quantum numbers. The common

methods effective in the weakly correlated case will have massive errors or fail to find solutions altogether. There are methods that are robust and reasonably accurate in this regime such as projected Hartree-Fock [2], which will be discussed later, and pair coupled cluster doubles [3] which is equivalent to exact diagonalization in the subspace of unbroken electron pairs in repulsive systems. While these methods do not fail for strong or weak correlation, they typically miss much of the weak correlation beyond mean field, so they are not accurate for all systems.

The third category is reasonably accurate for both weak and strongly correlated systems, but suffers from other shortcomings. These are correlated methods, such as coupled cluster, that build on a spin-symmetry broken mean field wavefunction. Breaking spin symmetry can significantly improve calculated energies and compensate for many of the failings of methods for strongly correlated systems. This is due to the spin-broken mean field wavefunction's ability to describe some of the properties of a strongly correlated system. For example, breaking S^2 symmetry, an independent particle wavefunction can produce an antiferromagnetic structure. However, these methods suffer from qualitatively poor wavefunctions, as they do not have the good quantum numbers that the true eigenstates must have. This means that calculated properties other than the energy may be poorly described.

There are certainly other methods addressing both of these problems such as DMRG, various forms of quantum Monte Carlo, and exact diagonalization, but they typically have significant limitations. These problems can include computational cost, limitations to particular systems, or lack of a 'black box' treatment meaning that they require careful tailoring to each individual problem.

In this work, I present a new basis for methods that work to solve some of the limitations of available methods. We will also explore some potential future endeavors

in order to select strong candidates for further development onto effective universal methods that address both types of correlation. We will explore a new application of some correlated wavefunctions by directly introducing correlation factors rather than as a set of excitations. We will also examine a set of methods combining wavefunction forms for weak and strong correlation methods as possible routes for building a better overall wavefunction.

Chapter 2

Background Material

The methods discussed in this work build on a number of long studied theories that require some review. Here we will discuss the two methods that prevalent in this work, namely Hartree-Fock and symmetry projection, as well as the quantum systems that we will be using to test the methods.

2.1 Notation

There are two forms of indexing used in this work to label the single particle orbitals defining the Hilbert space of the problems. For many of the equations the spatial and spin indices of the second-quantized operators will be represented separately (eg. $i\sigma$) with a lower-case index for the spatial coordinate summed over N orbitals, or sites, and a Greek symbol ($\sigma = \uparrow, \downarrow$) for the spin. There will be some cases where, due to length or form of the equations, this will be replaced with a single composite index represented by an upper-case letter (eg. I) summed over $2N$ spin-orbitals. This will be used to reduce the size and improve the readability of some of the longer results where explicitly identifying the spin sectors is less important.

There are a number of equations that will have indices representing either the basis of fermion operators with which the Hamiltonian is defined $\{a_I, a_J^\dagger\}$, sometimes called the lattice or Hamiltonian basis, and the basis of the quasiparticles for the Slater determinant wavefunction, sometimes called the canonical basis. This basis

consists of N_o occupied and N_v virtual states. In the index notation, indices O, o and V, v will be reserved to represent occupied and virtual state indices respectively. The corresponding Fermionic operators will be $\{c_O, c_O^\dagger, c_V, c_V^\dagger\}$. Other indices will represent summation in the Hamiltonian basis unless otherwise stated. These bases will be discussed further in Section 2.3.

2.2 Hamiltonians

In order to test the methods discussed here, it is important to have relevant systems for benchmark calculations. Ideally these systems should be simple to implement, but have nontrivial solutions. In addition it is useful if these models are exactly solvable or high quality numerical results are available as a reference. Here I introduce the main model Hamiltonian that will be used to test the methods discussed later as well as the Hamiltonian for molecular systems used in a later chapter.

2.2.1 The Hubbard Model

The nearest-neighbor, repulsive Hubbard Hamiltonian is a well studied, non-trivial system useful for testing new methods,

$$H = -t \sum_{\langle ij \rangle} (a_{i\uparrow}^\dagger a_{j\uparrow} + a_{i\downarrow}^\dagger a_{j\downarrow}) + U \sum_i n_{i\uparrow} n_{i\downarrow}, \quad (2.1)$$

where $a_{i\sigma}^\dagger$ and $a_{i\sigma}$ are on-site Fermionic creation and annihilation operators, and $n_{i\sigma} = a_{i\sigma}^\dagger a_{i\sigma}$ is the on-site number operator. The summation over $\langle ij \rangle$ indicates that we only include index pairs that are nearest neighbors. The hopping amplitude t represents the kinetic tunneling between neighboring sites, and the short-range Coulomb repulsion U penalizes any site that is doubly occupied. In the limit where

the on-site interaction U is zero, the Hamiltonian is solved by a simple product of plane wave states. However, when the interaction is nonzero, there is a competition between the plane wave solution and a Néel structure in the ground state. This competition produces nontrivial solutions and a rich mixture of both weak and strong correlation regimes depending on the value of $\frac{U}{t}$. If the lattice in consideration is one dimensional, there exists a closed form solution to this system, and the exact ground state energy can be calculated with the Lieb-Wu equations. In addition, the exact energy and other properties can be calculated for reasonably sized 1D systems using existing DMRG packages for the comparison of other properties. While there is not solution for a two dimensional lattice, extensive studies and benchmark calculations with highly accurate methods have been published providing a large amount of reference data for comparison. The simplicity of the system combined with the non-trivial solution and extensive studies makes the Hubbard model a useful benchmarking tool when studying the accuracy and effectiveness of new methods. All the energies reported for Hubbard systems will be given in units of t .

2.2.2 General Electronic Hamiltonians

In the case where we consider general electronic structure calculations, we will refer to the second quantized, non-relativistic Hamiltonian for molecular systems. The real-space form of this Hamiltonian with a convenient selection of units is written as,

$$H = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_A \frac{1}{2M_A} \nabla_A^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{i>j} \frac{1}{r_{ij}} + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}}, \quad (2.2)$$

where r_i represents the position of electron i , R_A represents the position of nucleus A , and M_A and Z_A are the nuclear charges and masses respectively. This Hamiltonian is

treated with the Born-Oppenheimer approximation, where we assume that the electron and nuclear degrees of freedom are separable. In this approximation we consider the nuclear positions to be fixed-points, or at least subject to slower, classical fluctuations compared to the much less massive electrons. This means we now only consider the electron positions as the quantum variables, and the effective Hamiltonian becomes,

$$H = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{i>j} \frac{1}{r_{ij}} + \text{const.} \quad (2.3)$$

This Hamiltonian is treated via projection into a set of basis functions, $|I\rangle$, in order to write it as a second-quantized operator,

$$H = \sum_{IJ} t_{IJ} a_I^\dagger a_J + \frac{1}{4} \sum_{IJKL} v_{IJKL} a_I^\dagger a_J^\dagger a_L a_K, \quad (2.4)$$

where a_I^\dagger and a_I are the electronic creation and annihilation operators for basis function $|I\rangle$ and,

$$t_{IJ} = \langle I | \left(-\frac{1}{2} \sum_i \nabla_i^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} \right) | J \rangle, \quad v_{IJKL} = \langle IJ | \left(\sum_{i>j} \frac{1}{r_{ij}} \right) | KL \rangle, \quad (2.5)$$

are the one and two-particle overlaps.

2.3 Hartree-Fock

The Hartree-Fock mean field wavefunction is a common starting point for the development of more advanced, correlated wavefunctions. The selection of an appropriate reference wavefunction is an important consideration in this work, so it is necessary to discuss the mean-field solution before proceeding further. This will also demonstrate some of the important concepts for calculating overlaps and expanding expectation

values of many-body operators.

The Hartree-Fock method minimizes the energy of a given system for a single Slater determinant wavefunction [4]. There are two main methods for calculation of the Hartree-Fock wavefunction. We start with a general, two-body Hamiltonian written in second quantization,

$$H = \sum_{IJ} t_{IJ} a_I^\dagger a_J + \frac{1}{4} \sum_{IJKL} v_{IJKL} a_I^\dagger a_J^\dagger a_L a_K. \quad (2.6)$$

The Hamiltonian is often initially written in some physically relevant basis such as lattice sites or atomic orbitals with the corresponding creation and annihilation operators a_I and a_I^\dagger . We then seek to minimize the energy of this Hamiltonian with a Slater determinant wavefunction with N_o particles in $2N$ states,

$$|\Phi\rangle = \prod_O c_O^\dagger |-\rangle, \quad c_O^\dagger = \sum_I C_{IO}^* a_I^\dagger, \quad (2.7)$$

where $|-\rangle$ is the bare vacuum. The operators c_J form a new basis defining the Slater determinant. The coefficients C are an $M \times N_o$ transformation matrix between the two bases and the columns are a subset of N_o occupied states in an orthonormal basis of M single-particle states. Expectation values are evaluated via the one-particle density matrix in the original basis,

$$\rho_{IJ} = \langle \Phi | a_J^\dagger a_I | \Phi \rangle = \sum_O C_{IO} C_{JO}^*, \quad (2.8)$$

where K is summed over the occupied orbitals. The expectation value of higher-body operators can be evaluated via Wick's theorem as an antisymmetric expansion of

one-particle density matrices. [5] For a two-body operator we get,

$$\langle \Phi | a_I^\dagger a_J^\dagger a_L a_K | \Phi \rangle = \rho_{KI} \rho_{LJ} - \rho_{LI} \rho_{KJ}. \quad (2.9)$$

From this result, the mean-field energy expression is straightforward to represent,

$$E = \sum_{IJ} t_{IJ} \rho_{JI} + \frac{1}{2} \sum_{IJKL} v_{IJKL} \rho_{KI} \rho_{LJ}. \quad (2.10)$$

Here we have used the requirement that the elements of v must obey Fermionic antisymmetry ($v_{IJLK} = -v_{IJKL}$).

Solving for the optimal Hartree-Fock wavefunction is typically done with one of two methods. Each method has certain advantages and we will discuss both. The first method uses the Hartree-Fock self-consistent field equations (SCF). These equations are constructed by finding the minimum in the energy under the normalization constraint $S = I$, where S is the $N_o \times N_o$ overlap matrix,

$$S_{OO'} = \sum_I C_{IO}^* C_{IO'} \quad (2.11)$$

If we search for a stationary point in the energy, under this constraint, we can construct the Hartree-Fock equations [5] by varying the energy with respect to C^* ,

$$\delta(E - \text{tr}[\epsilon(S - I)]) = \left(\frac{\partial E}{\partial \rho} \frac{\partial \rho}{\partial C^*} - \frac{\partial \text{tr}(\epsilon S)}{\partial C^*} \right) \delta C^*, \quad (2.12)$$

where ϵ is a set of Lagrange multipliers. The derivative of the energy with respect to

the density is referred to as the Fock matrix,

$$F_{IJ} = \frac{\partial E}{\partial \rho_{JI}} = t_{IJ} + \sum_{KL} v_{IKJL} \rho_{LK}. \quad (2.13)$$

As C^* is free to vary within the limits of the constraint, stationarity is enforced by requiring the term in parentheses on the right side of 2.12 to vanish,

$$\sum_{IJ} F_{IJ} \frac{\partial \rho_{JI}}{\partial C_{KO}^*} = \frac{\partial \text{tr}(\epsilon S)}{\partial C_{KO}^*}, \quad (2.14)$$

which simplifies to,

$$\sum_K F_{IK} C_{KO} = \sum_O C_{IO} \epsilon_{OO'} \quad (2.15)$$

Varying the energy with respect to C rather than C^* will get an equivalent condition ($C^\dagger F = \epsilon C^\dagger$).

We can modify this requirement into a more useful form. As the constraint is Hermitian, ϵ will also be Hermitian so we can decompose it as,

$$\epsilon_{OO'} = \sum_{K=1}^{N_o} U_{OK} \tilde{\epsilon}_K U_{O'K}^*, \quad (2.16)$$

where U is unitary. Inserting this into 2.15, multiplying on the right by U and defining $\bar{C} = CU$ results in,

$$\sum_K F_{IK} \bar{C}_{KO} = \bar{C}_{IO} \tilde{\epsilon}_O. \quad (2.17)$$

The rotation U does not affect any results as it only mixes the occupied orbitals amongst themselves. Expectation values with Slater determinants can be written

purely in terms of the density matrix which is invariant to occupied-occupied mixing,

$$\rho = \overline{CC^\dagger} = CUU^\dagger C^\dagger = CC^\dagger. \quad (2.18)$$

It is also invariant to mixing between the virtual states as the density matrix depends only on the occupied states.

Equation 2.17 tells us that the Hartree-Fock energy is at a stationary point when the density and Fock matrices share a common eigenbasis. This requirement is commonly written in the equivalent form,

$$[F, \rho] = 0, \quad (2.19)$$

as this can be easily calculated and tested for convergence. To see this, we write the eigenvalue decomposition of the two matrices when Equation 2.17 is satisfied,

$$F = D\mathcal{E}D^\dagger, \quad \rho = D\Lambda D^\dagger. \quad (2.20)$$

D is the full set of eigenvectors for F containing the coefficients for the occupied states C and the virtual states V ($D = [C \ V]$), \mathcal{E} is diagonal and contains the eigenvalues of F , and Λ is also diagonal containing ones for the occupied vectors C and zeros otherwise. As these matrices have the same eigenbasis, D , at convergence, the condition (2.19) will be satisfied as $D^\dagger D = I$ and \mathcal{E} and Λ commute since they are diagonal,

$$[F, \rho] = D[\mathcal{E}, \Lambda]D^\dagger = 0. \quad (2.21)$$

If the Hartree-Fock equations have not been satisfied, the Fock matrix will have a different eigenbasis than the density matrix used to construct it and they will not

commute. This equation is solved by taking an initial guess for the density and constructing the Fock matrix to calculate a set of eigenvectors for a new trial guess. This process is iterated until the convergence criterion is met. This method is fast and converges rapidly when the initial guess is close to the solution, but suffers from instabilities in some cases, such as when the initial guess is poor and for systems where there are near degeneracies.

The second method for solving Hartree-Fock is via the calculation of the energy gradient and direct minimization of the energy. This is more robust and will find solutions, or at least local minima, when the initial guess is poor and when the system contains near degeneracies or strong correlation. The tradeoff is an increased cost for evaluating the equations and typically slower convergence to the solution. Thouless's theorem [6] shows that we can transform one Slater determinant wavefunction into another via a one body rotation,

$$|\Phi'\rangle = e^K |\Phi\rangle, \quad K = \sum_{OV} \kappa_{OV}^* c_V^\dagger c_O, \quad (2.22)$$

where κ is an $N_o \times N_v$ matrix. The gradients method simply finds the solution by minimizing the energy with respect to this rotation K . We construct the optimization in terms of a unitary rotation ($\tilde{K} = K - K^\dagger$) in order to reduce the cost when calculating the gradients. This is because the orbital derivative can be written as a commutator, reducing both the size of the operators to be evaluated and the number of dummy indices to sum over.

The Hartree-Fock energy is stationary with respect to occupied-occupied and virtual-virtual state mixing, so only the occupied-virtual mixing is included. We

write the energy in terms of κ and take the derivative to get the orbital gradient,

$$E[\kappa] = \langle \Phi | e^{-\tilde{K}} H e^{\tilde{K}} | \Phi \rangle, \quad G_{OV} = \left. \frac{\partial E}{\partial \kappa_{OV}^*} \right|_{\kappa=0} = \langle \Phi | [H, c_V^\dagger c_O - c_O^\dagger c_V] | \Phi \rangle. \quad (2.23)$$

As the elements of the operator \tilde{K} do not commute amongst themselves, we take the derivative centered around $\kappa = 0$ in order to maintain a simple, low-cost definition of the gradient.

For simple models such as the Hubbard Hamiltonian, it is often more efficient to evaluate the gradient in the Hamiltonian basis $\{a^\dagger, a\}$ and convert to the Slater determinant basis $\{c^\dagger, c\}$ rather than transform the Hamiltonian into the canonical basis of the Slater determinant,

$$G_{OV} = \sum_{IJ} C_{IO}^* (G_{JI}^0 - G_{IJ}^{0*}) C_{JV}, \quad (2.24)$$

$$G_{IJ}^0 = \langle \Phi | [H, a_I^\dagger a_J] | \Phi \rangle. \quad (2.25)$$

At step i in the minimization, the orbital coefficients are updated via iterative steps of some size ϵ taken in the direction of the negative gradient at each point.

$$D_{i+1} = D_i e^{-\epsilon K_i}, \quad (2.26)$$

$$K_i = \begin{pmatrix} 0 & -G_i \\ G_i^\dagger & 0 \end{pmatrix}, \quad (2.27)$$

where D is the full set of orbitals as defined below Equation 2.20. Using the calculated gradient, the energy can then be minimized via steepest descent as above, or via more advanced techniques such as conjugate gradient or Newton-Rhapson.

Both the gradient and SCF methods converge when the Hamiltonian no longer

mixes the occupied and virtual sector of the single particle states. The two methods simply minimize the variance of the energy with respect to two different but equivalent sets of variables. The gradient method minimizes with respect to rotations of the reference and the SCF equations directly with respect to the coefficients in C .

2.4 Symmetry Breaking and Restoration

A simple but effective method to forming a more advanced, correlated wavefunction out of a single Slater determinant is a process known as symmetry breaking and restoration. [7] This maintains the mean-field cost of Hartree-Fock, albeit with the inclusion of an integration grid, and is very effective at describing strong correlation with minimal complications.

2.4.1 Symmetry Broken Hartree-Fock

Traditional Hartree-Fock, known as restricted Hartree-Fock (RHF), conserves the spin-symmetry of the Hamiltonian. This is implemented via requiring the orbital coefficients for the up (α) and down (β) canonical orbitals to be the same, and this can be most easily demonstrated in the form of the resulting density matrix for the Hartree-Fock Slater determinant wavefunction written in the Hamiltonian basis [8],

$$\rho = \begin{pmatrix} \rho_{\alpha\alpha} & \rho_{\alpha\beta} \\ \rho_{\alpha\beta}^\dagger & \rho_{\beta\beta} \end{pmatrix}. \quad (2.28)$$

We sort the spin-orbitals such that the index values $[1, N]$ represent spin up, and values $[N + 1, 2N]$ represent spin down. RHF is the case where there is no mixing between the up (α) and down (β) states, and the up and down spatial orbitals are the same, or mathematically $\rho_{\alpha\beta} = 0$ and $\rho_{\alpha\alpha} = \rho_{\beta\beta}$. In cases where there is strong

correlation, the energy may be improved if we are willing to sacrifice some of the symmetry of the wavefunction. Breaking S^2 symmetry results in an unrestricted Hartree-Fock (UHF) wavefunction. In this case the up and down orbitals are different ($\rho_{\alpha\alpha} \neq \rho_{\beta\beta}$). We can take it one step further and allow the up and down orbitals to mix. This breaks S_z symmetry as well and is characterized by $\rho_{\alpha\beta} \neq 0$. This is referred to as generalized Hartree-Fock (GHF).

In systems with strong correlation, RHF produces qualitatively poor results as seen in Fig. 2.1. Symmetry breaking of the mean-field wavefunction can significantly improve the quality of calculated energies. However, the incorrect quantum numbers of the wavefunction will result in a poor description of other properties aside from the energy. For example, UHF will produce qualitatively correct energies for strongly correlated Hubbard systems, but the spin-spin correlation function will have massively overestimated values. In order to capitalize on the improved energies from the symmetry broken state without sacrificing good quantum numbers, we can restore the symmetry of the deformed reference via projection operators.

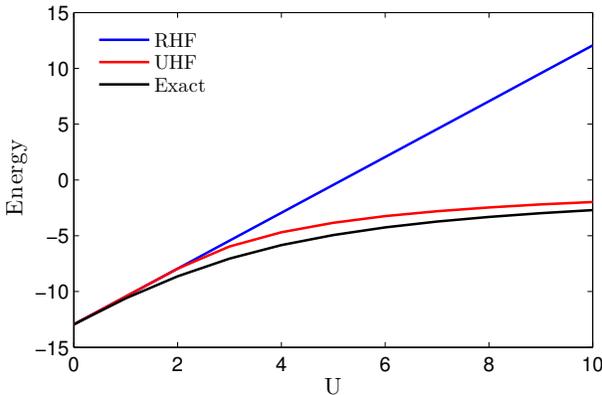


Figure 2.1 : Mean-field energies of a 10-site Hubbard ring as a function of the interaction strength.

2.4.2 Projected Hartree-Fock

The PHF wavefunction is constructed with a symmetry broken state, $|\Phi\rangle$ which is then projected into the symmetry adapted subspace with a projection operator P [2, 7],

$$|\Psi\rangle = P|\Phi\rangle. \quad (2.29)$$

The resulting wavefunction is no longer a single Slater determinant but a linear combination of nonorthogonal determinants. The energy of the projected wavefunction is then,

$$E = \frac{\langle\Phi|P^\dagger H P|\Phi\rangle}{\langle\Phi|P^\dagger P|\Phi\rangle} = \frac{\langle\Phi|H P|\Phi\rangle}{\langle\Phi|P|\Phi\rangle}. \quad (2.30)$$

Note here that the projection operator is idempotent ($P^2 = P$), is Hermitian for observable symmetries, and the projection commutes with operators that preserve the symmetry, such as the Hamiltonian. The operator P used in this work for spin projection is [7],

$$P_{mk}^s = |s; m\rangle\langle s; k|, \quad (2.31a)$$

$$P_{mk}^s = \frac{2s+1}{8\pi^2} \int d\Omega D_{mk}^{s*}(\Omega) R(\Omega), \quad (2.31b)$$

where $R(\Omega) = e^{i\alpha S_z} e^{i\beta S_y} e^{i\gamma S_z}$ and $D_{mk}^s(\Omega) = \langle s; m | R(\Omega) | s; k \rangle$ is the Wigner D-matrix. This operator projects a state on the right into a state with spin quantum numbers s and m by enforcing that the state is stationary with respect to spin rotations. In this work we will only consider the projection of a UHF state where S_z has a good quantum number m . This allows some simplification of the projection as the S_z rotations in $R(\Omega)$ will drop out. The spin projected UHF (SUHF) state generated by

the UHF state $|\Phi\rangle$ can then be represented as,

$$|SUHF\rangle = P_{mm}^s |\Phi\rangle = \frac{2s+1}{2} \int_0^\pi d\beta \sin \beta d_{mm}^s(\beta) e^{i\beta S_y} |\Phi\rangle, \quad (2.32)$$

where $d_{mm}^s(\beta) = \langle s; m | e^{i\beta S_y} | s; m \rangle$ is the Wigner d-matrix. Expectation values are then evaluated by Gaussian quadrature where each point of the integration is evaluated by transition density matrices between the rotated states. Details on SUHF expectation values can be found in Appendix A.

Chapter 3

The Lie Algebraic Similarity Transformation (LAST)

Similarity transformation methods have been applied in a number of contexts in condensed matter physics and electronic structure [9–14]. Jastrow-Gutzwiller correlation factors are also very popular as variational wave functions in quantum Monte Carlo and other applications [15–23]. Nonvariational solutions have also been discussed in the literature. Tsuneyuki [24] presented a Hilbert space Jastrow method based on a Gutzwiller factor and applied it to the one-dimensional (1D) Hubbard model, minimizing its energy variance as in the transcorrelated method [25–27]. Neuscamman et al. [28] proposed many-body Jastrow correlators, diagonal in the lattice basis, and truncated them to a subset of sites matching a given pattern; these authors compared projective solutions with those obtained stochastically via Monte Carlo.

Here, we consider Hamiltonian transformations of the form $e^{-J}He^J$ based on Hermitian correlators J built from general two-body products of on-site operators over the entire lattice. The transformations here are generated by density (charge), spin, and Gutzwiller factor correlators, including density-spin crossed terms. Similar Jastrow-type correlators have been extensively discussed in the literature but almost always in a variational context [18]. Our transformed Hamiltonian is non-Hermitian but can be solved in mean field via projective equations, similar in spirit to those of coupled cluster theory [1, 28]. In this sense, the model is an extension that fits under the generalized coupled cluster label [29–31]. The fundamental difference is

that traditional coupled cluster is formulated with particle-hole excitations out of a reference determinant via a non-Hermitian cluster operator; the present model is constructed with on-site Hermitian correlators.

The Hausdorff series resulting from the nonunitary similarity transformation $e^{-J}He^J$ can be analytically summed. This result follows from Lie algebraic arguments [32] after recognizing that both the Hamiltonian and the correlator J can be written in the basis of generators of an enveloping algebra built from on-site operators. Topologically, our transformation is noncompact and yields a non-Hermitian Hamiltonian, whereas traditional canonical transformations are almost always chosen to be unitary, thus compact, and preserve hermiticity. There is a mistaken belief that quantum canonical transformations must be unitary [33]; this is not correct even in the linear case [2]. From this perspective, traditional coupled cluster exponentiates the shifts of a nilpotent algebra, whose Hausdorff series truncates at the fourth commutator (for a two-body H). For two-body correlators, our model leads to a renormalized N -body Hamiltonian that produces locally weighted orbital rotations of a reference state, leading to expectation values between nonorthogonal determinants. Here, we introduce the main mathematical results in a self-contained manner, touching upon the physical aspects of the model, and present benchmark applications to the 1D and 2D Hubbard models.

3.1 Jastrow-type Correlation Factors

We seek to construct a theory similar in spirit to coupled cluster, but using correlation factors constructed of products of on-site spin-density operators $n_{i\sigma} = a_{i\sigma}^\dagger a_{i\sigma}$. These operators have many convenient properties when constructing a theory. They are idempotent, commute amongst themselves, and have a simple commutation relation

with the elemental fermion operators,

$$n_{i\sigma}^2 = n_{i\sigma}, \quad [n_{i\sigma}, n_{j\sigma'}] = 0, \quad [n_{i\sigma}, a_{j\sigma'}^\dagger] = \delta_{ij}\delta_{\sigma\sigma'}a_{j\sigma'}^\dagger. \quad (3.1)$$

We can use them to construct a two-body correlation factor,

$$J = \frac{1}{2} \sum_{i\sigma, j\sigma'} \alpha_{i\sigma, j\sigma'} n_{i\sigma} n_{j\sigma'}, \quad (3.2)$$

with α real, symmetric and containing zeros on the diagonal. This operator contains several types of correlations between all of the sites, or orbitals in the system. The main correlations of interest are, $N_i N_j$, $S_i^z S_j^z$, $N_i S_j^z$, and $D_i + D_j$, where

$$N_i = n_{i\uparrow} + n_{i\downarrow}, \quad S_i^z = n_{i\uparrow} - n_{i\downarrow}, \quad D_i = n_{i\uparrow} n_{i\downarrow}. \quad (3.3)$$

These are the local number, spin and double occupancy operators. The double occupancy term, $\sum_i n_{i\uparrow} n_{i\downarrow}$, is often referred to as the Gutzwiller factor as it was introduced in a similar capacity in earlier work by Martin Gutzwiller [20].

The correlation factor J is applied to a mean-field reference, $|\Phi\rangle$ via an exponential to produce a more correlated wavefunction, $|J\rangle = e^J |\Phi\rangle$. The exponential acts on the reference by increasing or suppressing the correlations between different orbitals in the wavefunction based on the values of the parameters in α . Wavefunctions of this form have been widely used in methods such as variational Monte Carlo. In order to produce a set of equations that can be evaluated analytically, the exponential can be applied as a similarity transformation of the Hamiltonian. By acting with the inverse of the exponential operator on Schrödinger's equation, we have a new, equivalent

condition for the exact solution,

$$He^J|\Psi\rangle = Ee^J|\Psi\rangle \quad \Rightarrow \quad e^{-J}He^J|\Psi\rangle = E|\Psi\rangle. \quad (3.4)$$

In cases where the similarity transformation can be evaluated, this provides a strong starting point to construct approximate methods. We refer to this as the Lie algebraic similarity transformation (LAST) as the correlation factors in J correspond to a set of Cartans in the algebra of all possible operators of a two-site system. Approximate methods are constructed by replacing the exact solution $|\Psi\rangle$ with an approximate wavefunction.

3.2 Evaluating the Similarity Transformation

The similarity transformation can be evaluated by looking at the transformation of an individual creation operator (Appendix B),

$$e^{-J}a_{i\sigma}^\dagger e^J = e^{-J_{i\sigma}}a_{i\sigma}^\dagger, \quad J_{i\sigma} = \sum_{j\sigma'} \alpha_{i\sigma,j\sigma'} n_{j\sigma'}. \quad (3.5)$$

The two-body similarity transformation becomes a site-dependent, one-body rotation that commutes with the transformed creation operator. Using this result and its conjugate, we can easily apply the transformation to one-body operators,

$$e^{-J}a_{i\sigma}^\dagger a_{j\sigma'} e^J = e^{-J}a_{i\sigma}^\dagger e^{-J}e^J a_{j\sigma'} e^J = e^{-J_{i\sigma}}a_{i\sigma}^\dagger a_{j\sigma'} e^{J_{j\sigma'}}. \quad (3.6)$$

The expectation value of this operator,

$$\langle \Phi | e^{-J_{i\sigma}} a_{i\sigma}^\dagger a_{j\sigma'} e^{J_{j\sigma'}} | \Phi \rangle, \quad (3.7)$$

can be evaluated by acting with the exponential term on the Slater determinant $|\Phi\rangle$,

$$\begin{aligned}
e^{J_{i\sigma}}|\Phi\rangle &= e^{J_{i\sigma}} \prod_{\sigma\gamma} (c_{\sigma\gamma}^\dagger)|-\rangle \\
&= \prod_{\sigma\gamma} \left(e^{J_{i\sigma}} c_{\sigma\gamma}^\dagger e^{-J_{i\sigma}} \right) |-\rangle \\
&= \prod_{\sigma\gamma} \left(\sum_{k\sigma'} C_{k\sigma',\sigma\gamma}^* e^{J_{i\sigma}} a_{k\sigma'}^\dagger e^{-J_{i\sigma}} \right) |-\rangle.
\end{aligned} \tag{3.8}$$

Using the result $[J_{i\sigma}, a_{k\sigma'}^\dagger] = \alpha_{i\sigma,k\sigma'} a_{k\sigma'}^\dagger$, this evaluates to,

$$e^{J_{i\sigma}}|\Phi\rangle = \prod_{\sigma\gamma} \left(\sum_{k\sigma'} C_{k\sigma',\sigma\gamma}^* \exp(\alpha_{i\sigma,k\sigma'}) a_{k\sigma'}^\dagger \right) |-\rangle, \tag{3.9}$$

$$= |\Phi_R^{i\sigma}\rangle. \tag{3.10}$$

We can see that this can be written as a Slater determinant,

$$|\Phi_R^{i\sigma}\rangle = \prod_{\sigma\gamma} r_{\sigma\gamma}^{i\sigma^\dagger} |-\rangle, \tag{3.11}$$

$$r_{\sigma\gamma}^{i\sigma^\dagger} = \sum_{k\sigma'} R_{k\sigma',\sigma\gamma}^{i\sigma} a_{k\sigma'}^\dagger, \tag{3.12}$$

$$R_{k\sigma',\sigma\gamma}^{i\sigma} = C_{k\sigma',\sigma\gamma}^* \exp(\alpha_{i\sigma,k\sigma'}). \tag{3.13}$$

The array $R^{i\sigma}$ is the transformation that defines the occupied orbitals of $|\Phi_R^{i\sigma}\rangle$, as C^* does for $|\Phi\rangle$. The superscript represents the dependence on the transformed operator in Equation 3.5. The transformation on the bra will produce the same result with the substitution $\alpha \rightarrow -\alpha$. Equation 3.7 can therefore be treated as a matrix element between two Slater determinants.

For the two Slater determinants,

$$|\Phi_A\rangle = \prod_{o\gamma} \left(\sum_{i\sigma} A_{i\sigma,o\gamma}^* a_{i\sigma}^\dagger \right) |-\rangle, \quad (3.14)$$

$$|\Phi_B\rangle = \prod_{o\gamma} \left(\sum_{i\sigma} B_{i\sigma,o\gamma}^* a_{i\sigma}^\dagger \right) |-\rangle, \quad (3.15)$$

under the condition that $\langle \Phi_A | \Phi_B \rangle \neq 0$, the matrix elements can be calculated as [5],

$$\langle \Phi_A | a_{j\sigma'}^\dagger a_{i\sigma} | \Phi_B \rangle = \det(S) \rho_{i\sigma,j\sigma'} \quad (3.16)$$

where,

$$S_{o\gamma,o'\gamma'} = \sum_{i\sigma} A_{i\sigma,o\gamma}^* B_{i\sigma,o'\gamma'}, \quad (3.17)$$

$$\rho_{i\sigma,j\sigma'} = \sum_{o'\gamma',o\gamma} B_{i\sigma,o'\gamma'} (S^{-1})_{o\gamma,o'\gamma'} A_{j\sigma',o\gamma}^*. \quad (3.18)$$

The density ρ reduces back to the original definition of the density matrix elements with the same Slater determinant on the right and left when $A = B$ and $S = I$ (Equation 2.8).

3.3 The Working Equations

As we cannot feasibly solve Eq. 3.4 exactly, we can take a similar approach to coupled cluster methods. By approximating the solution to be a Slater determinant wavefunction, we can construct a set of equations in order to find the parameters in α such that Schrödinger's equation is approximately satisfied. The transformed Hamiltonian, $\bar{H} = e^{-J} H e^J$, is non-Hermitian, so we cannot simply variationally minimize the energy. We would however like to approximate the variational energy

expression,

$$E_v = \frac{\langle \Phi | e^J H e^J | \Phi \rangle}{\langle \Phi | e^{2J} | \Phi \rangle}. \quad (3.19)$$

This energy can be written in terms of the transformed Hamiltonian,

$$E_v = \frac{\langle \Phi | e^{2J} \bar{H} | \Phi \rangle}{\langle \Phi | e^{2J} | \Phi \rangle}, \quad (3.20)$$

and in order to be tractable, we will make a linear approximation to the left state,

$$\langle \Phi | e^{2J} \approx \langle \Phi | (1 + \tilde{J}), \quad \tilde{J} = \frac{1}{2} \sum_{i\sigma, j\sigma'} z_{i\sigma, j\sigma'} n_{i\sigma} n_{j\sigma'}. \quad (3.21)$$

This results in the approximate energy expression,

$$E_J = \frac{\langle \Phi | (1 + \tilde{J}) \bar{H} | \Phi \rangle}{\langle \Phi | (1 + \tilde{J}) | \Phi \rangle}. \quad (3.22)$$

This energy expression is similar in spirit to coupled cluster, a different method that utilizes a Hamiltonian similarity transformation. [1] The difference is that \tilde{J} does not generate excitations orthogonal to the reference determinant, so we must include the normalization. As we approximate the left state, this energy is not bound by the variational principle. There are in fact cases where the values of α and z can be chosen to make the energy arbitrarily large and negative. We can still search for local minima by solving for stationary points,

$$\frac{\partial E_J}{\partial z_{i\sigma, j\sigma'}} = 0 \quad \Rightarrow \quad \langle \Phi | n_{i\sigma} n_{j\sigma'} \bar{H} | \Phi \rangle = E_J \langle \Phi | n_{i\sigma} n_{j\sigma'} | \Phi \rangle, \quad \forall i\sigma, j\sigma'. \quad (3.23)$$

This indicates that the amplitudes are selected to satisfy Schrödinger's equation in the subspace spanned by $\{\langle \Phi | n_{i\sigma} n_{j\sigma'}, \langle \Phi | \}$. In addition, it can be shown that the z

dependance in the energy will vanish under this condition and the energy will reduce to $E_J = \langle \Phi | \bar{H} | \Phi \rangle$.

The values in z are still required to calculate other properties however, and these can be obtained by requiring the energy to be stationary with respect to α as well,

$$0 = \frac{\partial E_J}{\partial \alpha_{i\sigma, j\sigma'}} = \frac{\langle \Phi | (1 + \tilde{J}) [\bar{H}, n_{i\sigma} n_{j\sigma'}] | \Phi \rangle}{\langle \Phi | (1 + \tilde{J}) | \Phi \rangle}, \quad \forall i\sigma, j\sigma'. \quad (3.24)$$

The expectation value of a given operator \mathcal{O} is then,

$$\langle \mathcal{O} \rangle_J = \frac{\langle \Phi | (1 + \tilde{J}) e^{-J} \mathcal{O} e^J | \Phi \rangle}{\langle \Phi | (1 + \tilde{J}) | \Phi \rangle}. \quad (3.25)$$

3.4 LAST on the Hubbard Model

In order to examine the effectiveness of this method, we will evaluate benchmark results for the Hubbard Hamiltonian. As a reminder, the Hubbard model Hamiltonian is,

$$H = -t \sum_{\langle ij \rangle \sigma} a_{i\sigma}^\dagger a_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (3.26)$$

This has a number of advantages when working with LAST. Number, charge, and double occupancy correlations all play a key role in the ground state, and the simple form of the Hamiltonian leads to relatively easy equations to evaluate. The transformed Hamiltonian becomes,

$$\bar{H} = -t \sum_{\langle ij \rangle \sigma} e^{-J_{i\sigma}} a_{i\sigma}^\dagger a_{j\sigma} e^{J_{j\sigma}} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (3.27)$$

The correlation factors adjust the one-body hopping term with local one-body rotations but leaves the on-site interaction unchanged as it commutes with J . This is not

the case in general as will be seen in chapter 5.

We present benchmark calculations for one- and two-dimensional Hubbard systems with a Hartree-Fock Slater determinant reference. Unless otherwise stated, the calculations include Gutzwiller, density-density, and spin-spin terms, with energy in units of t . The correlation energy is measured with respect to restricted Hartree-Fock (RHF) energies.

In Figure 3.1 we compare the correlation energy captured for eight-hole doped systems with periodic and non-periodic boundaries. The theory is most accurate for systems with few particles and open boundary conditions, but as we increase the size of the system, finite size effects are suppressed, and we begin to approach the thermodynamic limit while still recovering more than 95% of the correlation energy. We produce highly accurate results for doped systems and find some reduction in the quality as we approach the thermodynamic limit but still find significant improvements.

We have applied the method to a select set of two-dimensional Hubbard lattices where high quality reference data are available (Table 3.1). By screening the incorrect double occupancy with the Gutzwiller factor and incorporating corrections to the correlations in the RHF wave function with the spin and density terms, most of the correlation energy is recovered, dramatically improving the results. This supports the

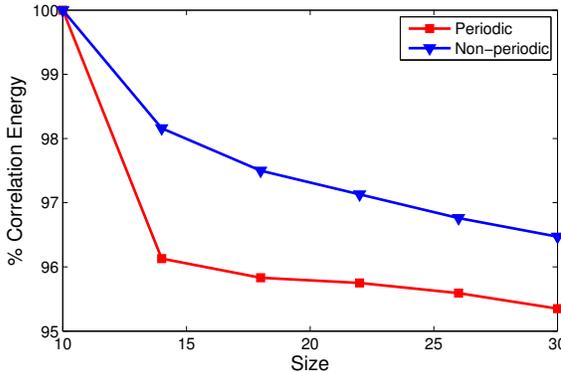


Figure 3.1 : Correlation energy of eight-hole doped Hubbard chains for $U = 2$ with open and closed boundaries on an RHF reference. Density matrix renormalization group (DMRG) is used to find exact energies for open systems [34, 35].

method as a cost-effective way to treat larger systems with high accuracy.

We can calculate other significant quantities using

Equation 3.25, and results agree well with other state-of-the-art methods. Figure 3.2 shows the discrete Fourier

transform of the spin-spin correlation function $S(i) = \langle S_0^z S_i^z \rangle$ for a one-dimensional Hubbard ring. We find that the Jastrow correlator adds most of the correct correlation on an otherwise smooth background. The function is only slightly underestimated at $q = \pi$, unlike the comparatively flat reference, and has the correct long-range decay.

If the wave function $|\Phi\rangle$ is a right eigenstate of the transformed Hamiltonian, then $e^J|\Phi\rangle$ is a solution to the original Hamiltonian, and we expect good approximations to $|\Phi\rangle$ and J to reflect this. In order to attest to the power of the method, we compare transformed and variational energies,

$$E_v = \frac{\langle \Phi | e^J H e^J | \Phi \rangle}{\langle \Phi | e^{2J} | \Phi \rangle}. \quad (3.28)$$

By directly computing overlaps with the exact wave function, we can determine

Size	N_o	U	E_{RHF}	E_J	E_{MC}	$\%E_c$
6×6	24	4	-1.0546	-1.1684	-1.1853	87.06
6×6	24	8	-0.6097	-0.9845	-1.0393	87.25
8×8	28	4	-1.0078	-1.0659	-1.0718	90.78
8×8	44	4	-1.0542	-1.1693	-1.1858	87.75

Table 3.1 : Energy per site and portion of the recovered correlation energy (E_c) for 2D, periodic lattices with occupancy N_o , an RHF reference and released-constraint Monte Carlo (E_{MC}) [36,37] as the best estimate for the exact result.

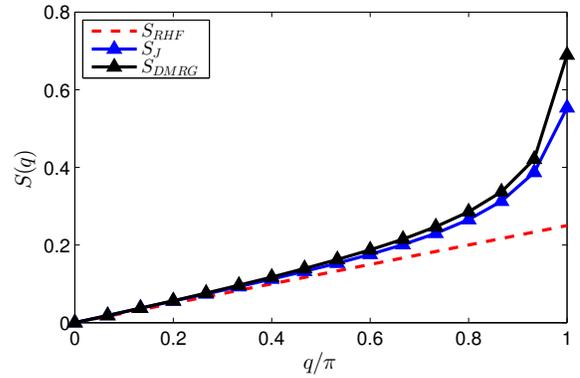


Figure 3.2 : The spin-spin correlation function in Fourier space calculated using Equation 3.25 (S_J) for a 30-site Hubbard ring at half filling and $U = 3$ compared to the RHF reference and DMRG. [34,35].

U	E_J	E_v	E_{exact}	E_{RHF}	$ \langle 0 \Phi\rangle $	$ \langle 0 J\rangle $
1	-14.6983	-14.7003	-14.7147	-14.4758	0.9721	0.9972
2	-11.8486	-11.8765	-11.9543	-10.9758	0.8780	0.9815
3	-9.3925	-9.5059	-9.7488	-7.4758	0.7100	0.9378
4	-7.4688	-7.4745	-8.0883	-3.9758	0.5296	0.8711
5	-5.5017	-6.0807	-6.8531	-0.4758	0.3967	0.8544
6	-3.7983	-4.8766	-5.9165	3.0242	0.3086	0.8437

Table 3.2 : Energies and overlaps for the exact $|0\rangle$, RHF $|\Phi\rangle$, and correlated wavefunctions $|J\rangle$ for a 14-site ring, where $|J\rangle = e^J|\Phi\rangle/|\langle\Phi|e^{2J}|\Phi\rangle|^{\frac{1}{2}}$.

how close the correlated wave function is to the exact solution. We show results for a 14-site system in table 3.2 where the transformed (E_J) and variational (E_v) energies have both been calculated with the same correlation factors α that satisfy equation 3.23. There is strong agreement between the two energies in the weak-coupling regime, where the results are of excellent quality, and reasonable agreement at larger interaction strengths. This is further supported by the overlap of the reference and correlated wave functions with the true ground state. As $e^J|\Phi\rangle$ is close to the true ground state, the Schrödinger equation is nearly satisfied and the energy evaluation using the transformed Hamiltonian is close to the corresponding variational energy.

For the treatment of larger systems, the cost can be moderated by restricting the correlation amplitudes to include only local interactions. For sufficiently weak U , the correlations can be limited to short range without significant impact on the quality of the results. Figure 3.3 illustrates this effect. As is clear from the plot, weaker interactions benefit little from correlation beyond second-nearest neighbors. Truncation at range R results in $O(MR)$ equations ($O(R)$ for translationally invariant systems) instead of $O(M^2)$, greatly reducing the computational effort required. Additionally,

the cost for construction, inversion, and the determinant evaluation of the overlap matrix in Equation 3.16 can be reduced by a factor of M via an update of the overlap for each new iteration due to the simple diagonal structure of the transformations. Truncating the range of the transformation in this manner will restrict the range at which correlation functions calculated with Equation 3.25 will be accurate, and we believe this approximation is most appropriate in systems where correlations decay rapidly.

There is some reduction in accuracy near half filling. This can be addressed using a spin-broken reference (Table 3.3). Whereas the RHF reference has large ionic contributions (zero or double occupancies), the UHF wave function possesses the correct qualitative antiferromagnetic character near half filling. As all two-body on-

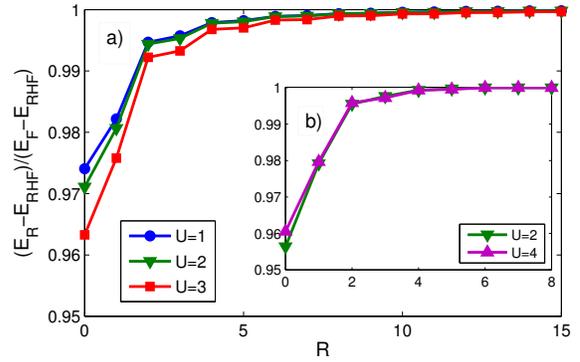


Figure 3.3 : The cumulative fraction of correlation energy captured by limiting the range R of the correlators compared to the full set (E_F) for a 70-site ring (a), and 8×8 2D lattice (b), at half filling with an RHF reference. $R = 0$ includes the Gutzwiller factor.

N_o	U	E_{RHF}	E_{UHF}	E_{RJ}	E_{UJ}	E_{ED}
14	2	-1.1172	-1.1644	-1.1634	-1.1920	-1.1982
14	4	-0.7344	-0.8808	-0.9018	-0.9595	-0.9840
14	8	0.0313	-0.5921	-0.5354	-0.6691	-0.7418
16	2	-1.0000	-1.0973	-1.0509	-1.1188	-1.1261
16	4	-0.5000	-0.7854	-0.6931	-0.8270	-0.8514
16	8	0.5000	-0.4619	-0.2235	-0.4873	-0.5293

Table 3.3 : Energies for 4×4 Hubbard lattices with RHF (E_{RJ}) and UHF (E_{UJ}) references including spin-density correlators ($S_i^z N_j$), compared to exact energies (E_{ED}) [36, 38].

site correlators are included in our model, it is sufficiently flexible to accommodate the necessary correlations depending on the nature of the reference. In the case of RHF, the largest contribution to the correlation energy is typically due to the Gutzwiller factor. Unlike the RHF case, we find a nonzero contribution from spin-density cross terms with a symmetry-broken reference as the up and down orbitals are no longer identical. Results improve significantly with the UHF reference, particularly for large values of U . However, for large U , we only recover a small fraction of the correlation energy with respect to UHF. Additional results are available in Appendix B.

From these initial calculations with the Jastrow correlation factors, we can see that this is an effective way to add some of the missing correlations to mean field wavefunctions. These contributions not only improve the energy, but can introduce the correct structure of the corresponding correlation functions into the system. However we see that we typically get around 90-95% of the missing correlation energy, and the calculation can become unstable for strongly correlated systems with an RHF reference. In addition, this method has $\mathcal{O}(M^4)$ cost on Hubbard when using the full set of correlators. This is not so bad, but for other systems where the two-body component is more complicated, this scaling will be increased up to $\mathcal{O}(M^7)$. In the next chapter we will simultaneously address both of these by reducing the problem, lowering the cost of evaluation, and improving the reference determinant to increase the accuracy.

Chapter 4

Improving the Reference Wavefunction

We have seen that Jastrow factors can serve to improve a mean-field wavefunction by introducing two-body correlations into a simple independent particle state. These mean-field references were selected as the solutions to the Hartree-Fock equations, meaning that they are the best possible independent particle wavefunction for the given system. However, in the presence of these correlation factors, there may be better wavefunctions on the mean-field level to accurately describe the true wavefunction. This can occur since specific types of incorrect correlation in the reference can be effectively screened by the correlators allowing greater flexibility on the wavefunction to accurately describe properties that the correlators are unable to capture. As an example, RHF on the Hubbard model exactly describes the hopping when there is no on-site interaction. When the repulsive term is turned on, the solution breaks symmetry in order to reduce the energy penalty from the large double occupancy, but in doing so the inter-site hopping is no longer accurately represented. If the orbitals are optimized in the presence of this on-site screening, the energy penalty can be reduced and the reference wavefunction is free to better represent the hopping in the final solution.

This optimization is done as in Hartree-Fock (Section 2.3) by requiring the energy to be stationary with respect to the variations in the reference determinant. The primary difference is that the similarity transformation results in a non-hermitian Hamiltonian, so we are not solving for a global minimum, and there is no condition

that the best possible left and right states are equivalent. As a result we adopt a biorthogonal ansatz for increased flexibility, and have found that the calculation is more stable than when the left and right states are forced to be equivalent. In order to keep the system simple and avoid increasing the cost of calculations, we will limit the correlators to only the on-site Gutzwiller factor. We have shown this to be the most important term in the previous chapter, and it makes the form of the orbital optimization equations particularly simple. In addition we will explore the further improvement of the results for strong and intermediate correlation by adding symmetry breaking and restoration on top of the orbital optimization.

As we are now limiting the correlaton factor to only contain the on-site Gutzwiller term ($J = \sum_i \alpha_i n_{i\uparrow} n_{i\downarrow}$) the similarity transformation of the Hamiltonian is simpler,

$$\bar{H} = -t \sum_{\langle i,j \rangle} \left((1 - \bar{g}_i n_{i\downarrow}) a_{i\uparrow}^\dagger a_{j\uparrow} (1 - g_j n_{j\downarrow}) + (1 - \bar{g}_i n_{i\downarrow}) a_{i\uparrow}^\dagger a_{j\uparrow} (1 - g_j n_{j\downarrow}) \right) + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (4.1)$$

We have used the fact that $n_{i\sigma}$ is an idempotent operator to rewrite the one-body exponents of the transformation (Equation 3.5) as a one-body operator,

$$e^{\alpha_i n_{i\sigma}} = 1 + g_i n_{i\sigma}, \quad g_i = e^{\alpha_i} - 1, \quad \bar{g}_i = e^{-\alpha_i} - 1. \quad (4.2)$$

This no longer requires the calculation of a site-dependant density matrix, so the evaluation complexity does not scale worse than the untransformed Hamiltonian.

4.1 The Calculation Scheme

We now describe the procedure to calculate the parameters in the ansatz. We refer to the following equations, before any symmetry projection operators are applied, as

the unrestricted Gutzwiller similarity transformation (UGST). In order to perform calculations, we construct a system of equations to solve for the amplitudes α_i and the reference Slater determinant. Since the transformed Hamiltonian is non-Hermitian, we do not expect the left and right eigenstates to be the same. We therefore use a biorthogonal ansatz for increased flexibility in the optimization,

$$E = \frac{\langle \Phi_L | (1 + \tilde{J}) \bar{H} | \Phi_R \rangle}{\langle \Phi_L | (1 + \tilde{J}) | \Phi_R \rangle}, \quad (4.3)$$

where $|\Phi_L\rangle$ and $|\Phi_R\rangle$ are UHF Slater determinants with different single particle bases and intermediate normalization $\langle \Phi_L | \Phi_R \rangle = 1$. We find this has some advantages aside from providing a more general ansatz than a single determinant. The results can be more accurate, particularly for doped lattices, and the stability and convergence rate of the reference optimization process is significantly improved.

The optimization conditions of all the degrees of freedom are now defined by taking derivatives of E as before,

$$\frac{\partial E}{\partial z_i} \propto \langle \Phi_L | n_{i\uparrow} n_{i\downarrow} (\bar{H} - E) | \Phi_R \rangle = 0, \quad \forall i, \quad (4.4)$$

$$\frac{\partial E}{\partial \alpha_i} \propto \langle \Phi_L | (1 + \tilde{J}) [\bar{H}, n_{i\uparrow} n_{i\downarrow}] | \Phi_R \rangle = 0, \quad \forall i. \quad (4.5)$$

This system of equations is solved for the parameters α_i and z_i .

4.1.1 Orbital Optimization

Up to this point this is all familiar from the previous chapter, but we need to have a way to select the optimal left and right Slater determinants. In order to optimize the right and left reference determinants, we use Hartree-Fock self-consistent field

equations as outlined in section 2.3. A generalized Fock matrix is constructed as in Section 2.3 as a derivative of GST energy with respect to the one-particle transition density ρ ,

$$F_{i\sigma,j\sigma'} = \frac{\partial E}{\partial \rho_{j\sigma',i\sigma}}, \quad (4.6)$$

where,

$$\rho_{i\sigma,j\sigma'} = \langle \Phi_L | c_{j\sigma'}^\dagger c_{i\sigma} | \Phi_R \rangle, \quad \rho = C_R C_L^\dagger.$$

We use the normalization condition,

$$C_L^\dagger C_R = I. \quad (4.7)$$

C_L and C_R are $M \times N_o$ matrices containing the occupied orbital coefficients of the left and right Slater determinants respectively, where M is the number of spin orbitals and N_o is the number of occupied states. These are the same as the orbital coefficients for Hartree-Fock in Section 2.3, but they describe different orbital bases which are not necessarily orthonormal. While the right and left states are no longer the same and the Hamiltonian is not Hermitian, the condition for a stationary point follows the same steps and will be the same result as the Hermitian case. The reference determinants are calculated with standard self-consistent Hartree-Fock iterations until F and ρ share common left and right eigenbases indicating we have reached a stationary point. [4] Both the amplitude equations and the Generalized Fock matrix have low computational cost, scaling as $\mathcal{O}(M^2)$ for the Hubbard Hamiltonian after construction of the density matrix.*

*Expressions for the energy and residuals are provided in Appendix. B

4.1.2 Using a PHF Reference

Once we have the reference determinants, we use the projection operators to restore symmetry of the wavefunction and further improve the results. We will refer to this as spin-projected UGST (SUGST). In principle, the reference optimization above can be done in the presence of the projection operators, but we find that this does not significantly change the results. In addition, the cost and difficulty of converging the equations is dramatically increased. As a result, we can simply choose to leave the reference determinants unchanged at this point and solve for a new set of amplitudes α^s and z^s in the presence of the projection. The expression for the energy and amplitude equations for the projected wavefunctions are the same as before,

$$E^s = \frac{\langle P_L^s | (1 + \tilde{J}^s) \overline{H}^s | P_R^s \rangle}{\langle P_L^s | (1 + \tilde{J}^s) | P_R^s \rangle}, \quad (4.8)$$

$$0 = \langle P_L^s | n_{i\uparrow} n_{i\downarrow} (\overline{H}^s - E^s) | P_R^s \rangle, \quad (4.9)$$

$$0 = \langle P_L^s | (1 + \tilde{J}^s) [\overline{H}^s, n_{I\uparrow} n_{I\downarrow}] | P_R^s \rangle \quad (4.10)$$

where \overline{H}^s and \tilde{J}^s are the transformed Hamiltonian and response operator as before but evaluated with α^s and z^s , and

$$\begin{aligned} |P_R^s\rangle &= \frac{P^s |\Phi_R\rangle}{\sqrt{\langle \Phi_L | P^s | \Phi_R \rangle}}, \\ \langle P_L^s | &= \frac{\langle \Phi_L | P^s}{\sqrt{\langle \Phi_L | P^s | \Phi_R \rangle}}, \end{aligned} \quad (4.11)$$

are the projected wavefunctions discussed in Section 2.4. As before, the values of α_i^s and z_i^s are calculated by requiring Equations 4.9 and 4.10 to be satisfied.

Expectation values of observables other than the energy are evaluated as before by

simply inserting the operator in question in place of the Hamiltonian in Equation 4.8.

$$\langle \mathcal{O} \rangle_J^s = \frac{\langle P_L^s | (1 + \tilde{J}^s) e^{-J^s} \mathcal{O} e^{J^s} | P_R^s \rangle}{\langle P_L^s | (1 + \tilde{J}^s) | P_R^s \rangle} \quad (4.12)$$

Operators such as spin, density, and double occupancy that commute with the transformation are still modified by the correlations when calculating their expectation value. It is important to note here that any operators which do not commute with the spin operators will require full spin integration with the full projection operator shown in Equation 2.31.

The iterative procedure for optimization of the reference determinants and correlation amplitudes is as follows:

1. Make an initial guess for C_L and C_R .
2. Solve Equation 4.4 for the amplitudes α .
3. Solve Equation 4.5 for the response amplitudes z .
4. Construct and diagonalize the Fock matrix (Equation 4.6) to build a new set of coefficients C_L and C_R .
5. Iterate 2-4 until the equations converge and the Hartree-Fock condition $[F, \rho] = 0$ is satisfied.
6. Solve Equation 4.9 for the amplitudes α^s .
7. Solve Equation 4.10 for the response amplitudes z^s .

The optimized parameters are then used to calculate energy and other properties with Eq. (4.3,4.8,4.12).

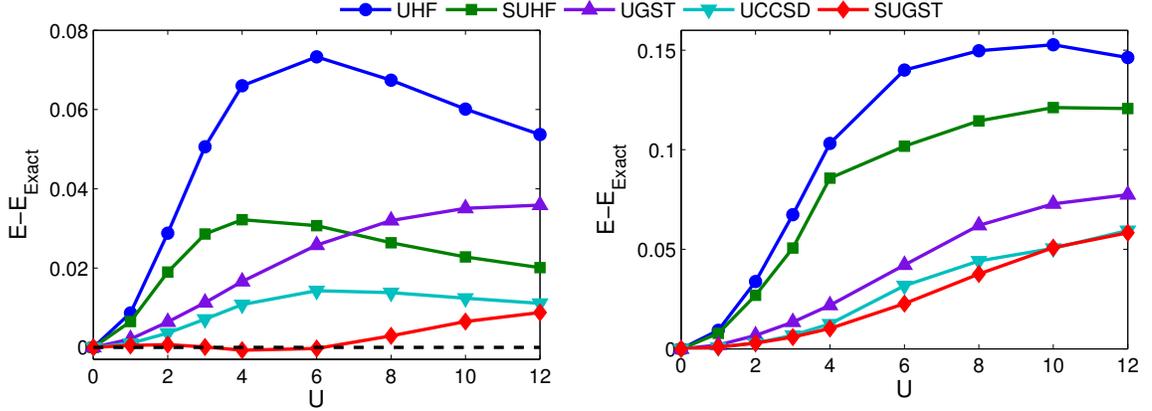


Figure 4.1 : Error in the energy per site for 4×4 square Hubbard lattices with 16 particles (left) and 14 particles (right) compared to exact diagonalization from Ref. [38].

4.2 Results and Discussions

We present benchmark calculations on Hubbard systems and compare the results to available accurate data. All calculations are performed on lattices with periodic boundary conditions, the spin state $s = m = 0$, and energies reported in units of t . We also compare to unrestricted coupled-cluster singles and doubles (UCCSD) where correlations are introduced with a Hamiltonian similarity transformation consisting of all single and double excitation operators that preserve S_z symmetry evaluated with the UHF determinant. [1] Tables are provided in Appendix B for direct comparison.

In Figure 4.1, we compare results for some 4×4 square lattices where exact energies are available. It is clear that SUHF and UGST have significant improvements over the UHF reference energy. When the two methods are combined in SUGST, the result is cumulative and we capture more of the correlation energy, typically more than UCCSD. The SUGST correlation is less than the sum of the SUHF and UGST correlation energy indicating there may be some overlap in the correlation energy recovered.

The quality of the results diminishes slightly for the doped systems as spin projection of a GHF state is likely better suited, but the energies remain similar to UCCSD.

A significant advantage of SUGST is the low cost of the calculations. As discussed above, UGST scales as $\mathcal{O}(N^3)$ in the number of sites which matches the observed times very closely (Figure 4.2).

The SUGST calculation formally scales as $\mathcal{O}(N^3 N_g)$ where N_g is the size of the integration grid, which is slightly lower than the observed rate. This may vary for different systems as the convergence rates can change. The low scaling means we can easily perform calculations on large systems with relatively little computational effort.

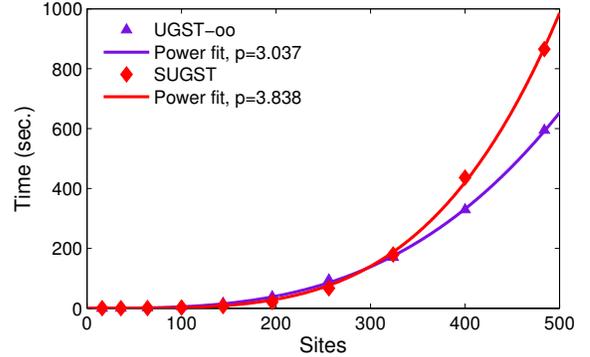


Figure 4.2 : Time required for the UGST orbital optimization and SUGST amplitude optimization with integration grid size equal to $\sqrt{N_{sites}}$ on half-filled Hubbard square lattices with $U = 4$

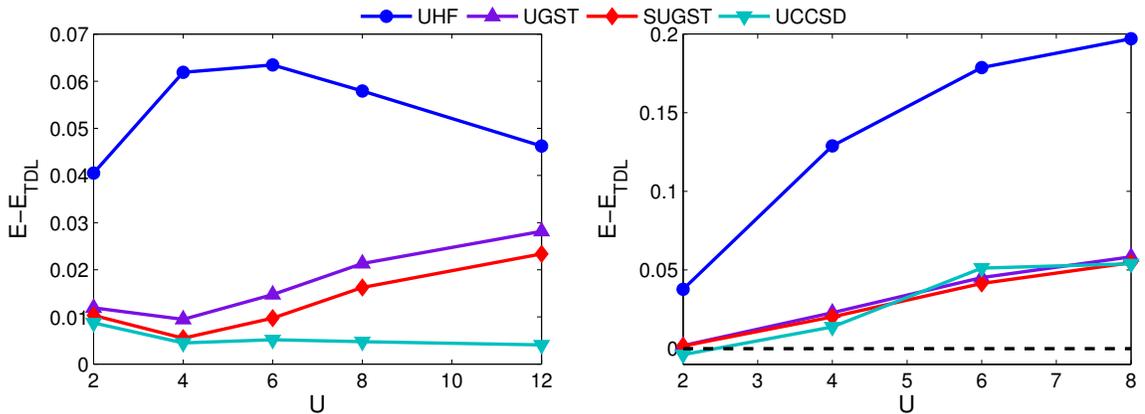


Figure 4.3 : Energy errors and finite-size effects per site for 10×10 square Hubbard lattices with $\langle n \rangle = 1$ (left) and $\langle n \rangle = 0.8$ (right) compared to UCCSD and averages of high quality thermodynamic limit calculations (E_{TDL}) from Refs. [39, 40] as accurate results for finite systems are limited.

We now apply the method to a set of larger Hubbard lattices at varying values of U and compare the results to UCCSD (Figure 4.3). There are some finite-size effects apparent for smaller values of U as the reference energies are borrowed from calculations for infinite systems. [39] Again we see significant improvement over mean-field when the transformation is applied and evaluated with the projected wavefunctions. Unlike the 4×4 case, there is some reduction in accuracy for larger U at half-filling. For the smaller systems, much of the correlation energy in this case was recovered through the projected wavefunction, and comparatively less was recovered by UGST than in the doped cases. In the larger systems, there is significantly less correlation energy per-site recovered with the projection, hence the larger errors.

We can see the reason for the reduction in quality as we increase the lattice size by observing the effects in the thermodynamic limit. As the computational cost of SUGST is low, it is a simple matter to perform calculations on very large systems. In Figure 4.4 we show the size effects on the energy per site of square Hubbard lattices. It is clear that SUGST

suffers from the same lack of size extensivity observed in PHF. [7] UGST converges to a thermodynamic limit as UHF does, but the additional correlation energy from the projection decreases as the system size increases once the thermodynamic limit is reached and eventually returns to the UGST energy per particle. There is a size intensive term in projection that yields a finite constant to be

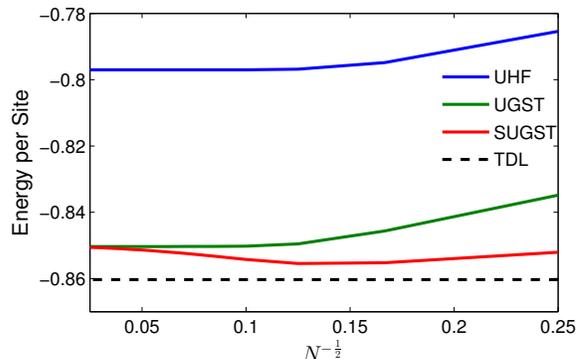


Figure 4.4 : Energies per site for the half-filled Hubbard model approaching the thermodynamic limit for N -site square lattices at $U = 4$, with an average of high quality results for an infinite system (TDL) from Ref. [39].

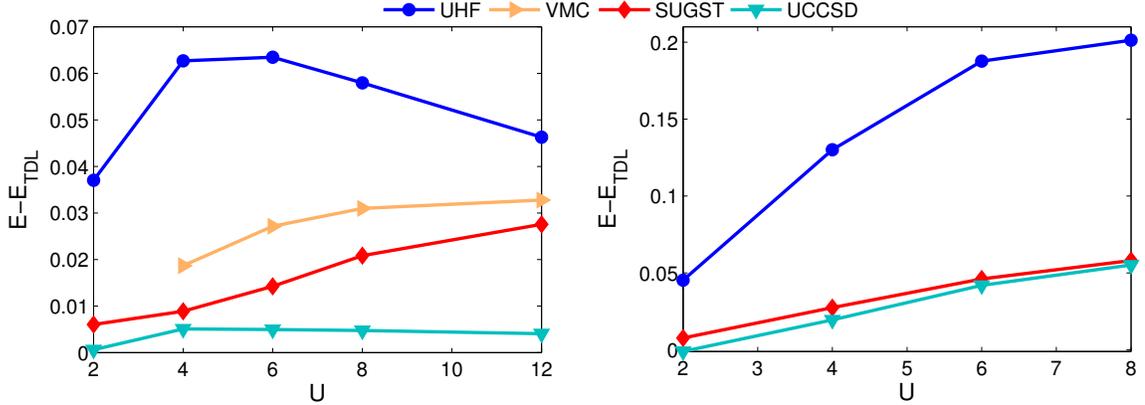


Figure 4.5 : Energy errors per site for 30×30 square Hubbard lattices with $\langle n \rangle = 1$ (left) and $\langle n \rangle = 0.8$ (right) compared to extrapolated variational Monte Carlo (VMC) from Ref. [41] as well as extrapolated UCCSD and averages of high quality thermodynamic limit calculations (E_{TDL}) from Ref. [39].

added to the infinite energy of an infinite system. [7]

If we now compare results in the thermodynamic limit for different values of U , we can see the previously observed behavior is maintained for large systems (Figure 4.5). As the SUHF wavefunction brings effectively no correlation energy per site for such large systems, we again see a reduction in accuracy for the strongly correlated case at half-filling with large U . We still find reasonable accuracy for $U = 4$ and the doped cases with results very close to the largest UCCSD lattices available. We can also compare the double occupancy

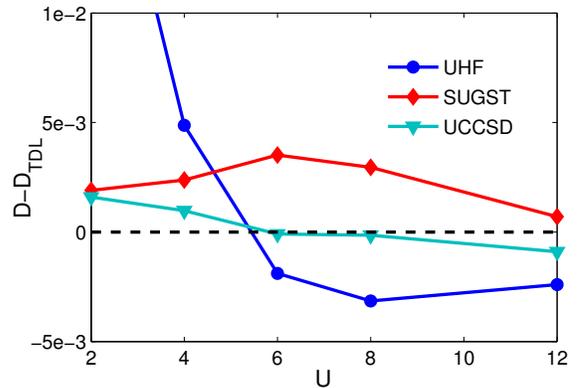


Figure 4.6 : Double occupancy errors per site for 30×30 square Hubbard lattices with $\langle n \rangle = 1$ compared to UCCSD and average of high quality thermodynamic limit calculations (D_{TDL}) from Ref. [39].

response densities. (Figure 4.6). While SUGST slightly overestimates the double occupancy, the error does not vary widely as UHF does.

We can also compare to the results taken from the literature [41] using variational Monte Carlo calculations with a Gutzwiller factor and an antiferromagnetic, mean-field reference. As the symmetry projection provides negligible improvement when approaching the thermodynamic limit, and all the local correlation factors equal a constant value for half-filling, this is a reasonable comparison with the variational solution of the wavefunction. In Figure 4.5, we can see that both the variational and projective methods have similar errors. We can further directly compare the wavefunctions by looking at the correlation amplitudes and antiferromagnetic order parameter in Table 4.1. While we do not get exactly the same energies and parameters, the results are very similar. We do not expect the results to be identical as the Gutzwiller wavefunction is not an exact solution. However, the similarity in the results indicates that we have made a good approximation to the variational solution

without the need for Monte Carlo sampling.

As the Gutzwiller factor only includes on-site terms, it provides significant improvement for short-range quantities such as the energy and double occupancy discussed above. In this method, longer-range correlations are left entirely to the reference wavefunc-

U	4	6	8	12
g_{VMC}^a	0.65	0.55	0.50	0.40
g_{SUGST}	0.6167	0.5205	0.4679	0.4147
$M_{VMC}^{a,b}$	0.58(2)	0.77(1)	0.86(1)	0.92(1)
M_{SUGST}	0.5851	0.7422	0.8352	0.9256

^a Results taken from Ref. [41].

^b Uncertainty for the last digit is given in parentheses.

Table 4.1 : Gutzwiller correlation factors (g) and antiferromagnetic order parameters (M) from SUGST on a 30×30 , half-filled lattice and variational Monte Carlo (VMC) extrapolated to the thermodynamic limit.

tion. We see that the spin-spin correlation function (Figure 4.7) quickly decays to a constant value and does not capture the correct long range decay of the exact correlation. There is significant improvement over the Hartree-Fock and projected results as the correlations are not severely over or underestimated in the medium range. The short range interactions are effectively screened, allowing the reference wavefunction to approximate the long-range effects more accurately within the limitations of the ansatz. As shown in Chapter 3, long-range effects can be correctly calculated if the Jastrow factor used for the transformation contains long-range terms.

It is clear that this method is an effective way of calculating the energy and short range correlations in the Hubbard lattice model. Including the projected wavefunction solves some of the failures of LAST in the strongly correlated limit. Limiting the correlation factor to double occupancy screening still results in accurate results and produces a theory that is only mean-field scaling.

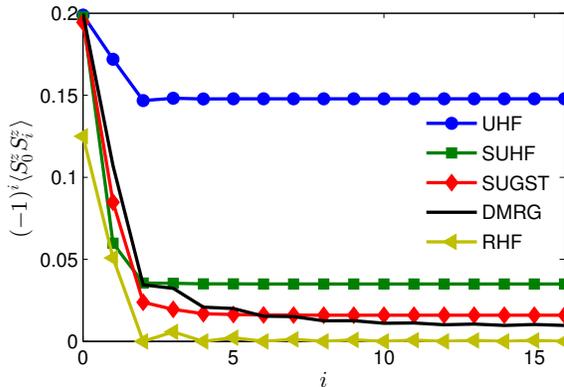


Figure 4.7 : Spin-spin correlation function with alternating sign for a 30-site Hubbard ring with $U = 4$ compared to the exact density matrix renormalization group theory (DMRG) result. [34, 35]

Chapter 5

LAST for General Hamiltonians

There is a natural question as to how well LAST performs when applied to general systems of the form,

$$H = \sum_{IJ} t_{IJ} a_I^\dagger a_J + \frac{1}{4} \sum_{IJKL} v_{IJKL} a_I^\dagger a_J^\dagger a_L a_K. \quad (5.1)$$

Applying LAST to this general Hamiltonian produces a new, correlated Hamiltonian,

$$\bar{H} = \sum_{IJ} t_{IJ} e^{-J_I} a_I^\dagger a_J e^{J_J} + \frac{1}{4} \sum_{IJKL} e^{\alpha_{IJ} - \alpha_{KL}} v_{IJKL} e^{-J_{IJ}} a_I^\dagger a_J^\dagger a_L a_K e^{J_{KL}}, \quad (5.2)$$

where I have used the notation $J_{IJ} = J_I + J_J$. The values in α can then be obtained as before by requiring,

$$0 = \langle n_I n_J (\bar{H} - \langle \bar{H} \rangle) \rangle. \quad (5.3)$$

There is however a very significant limitation. Evaluation of the two-body term requires the calculation of a transition density at cost $\mathcal{O}(N^3)$ for each term in the summation resulting in a cost of $\mathcal{O}(N^7)$ for the evaluation of the energy and the amplitude equations.

5.1 The Gutzwiller Similarity Transformed Hamiltonian

In order to simplify the problem and avoid some potential instabilities as before, we will examine this problem in the simple case where we limit the correlator to the

on-site Gutzwiller factor. $J = \sum_i \alpha_i n_{i\uparrow} n_{i\downarrow}$. By capitalizing on the simple form of the transformed Fermionic operators from Chapter 4, the transformed Hamiltonian takes the form,

$$\begin{aligned} \bar{H} &= \sum_{IJ} t_{IJ} (1 + \bar{g}_I n_{\bar{I}}) c_I^\dagger c_J (1 + g_J n_{\bar{J}}) \\ &+ \frac{1}{4} \sum_{IJKL} v_{IJKL} (1 + \bar{g}_I n_{\bar{I}}) c_I^\dagger (1 + \bar{g}_J n_{\bar{J}}) c_J^\dagger c_L (1 + g_L n_{\bar{L}}) c_K (1 + g_K n_{\bar{K}}). \end{aligned} \quad (5.4)$$

The bar notation indicates that I and \bar{I} have the same spatial coordinate, but opposite spin, and $g_I = e^{\alpha_I} - 1$, $\bar{g}_I = e^{-\alpha_I} - 1$ as before. Note that these are still only spatially dependent ($g_{i\uparrow} = g_{i\downarrow} = g_I$). This is a six-body operator, and evaluating the amplitude equations, requires the expectation value of an eight-body operator. As this results in an expansion via Wick's theorem with more than 40,000 terms, some simplification is in order.

If we restrict the reference to be RHF or UHF, we can separate the eight-body operator by spin blocks and evaluate it as separate expectation values of a three and five-body operator. This significantly reduces the number of terms in the expansion. This is most easily represented by explicitly writing all of the nonzero spin sectors of the Hamiltonian,

$$\begin{aligned} H &= \sum_{ij} \left(t_{ij}^{\alpha\alpha} a_{j\uparrow}^\dagger a_{i\uparrow} + t_{ij}^{\beta\beta} a_{j\downarrow}^\dagger a_{i\downarrow} \right) + \sum_{ijkl} v_{ijkl}^{\alpha\beta\alpha\beta} a_{i\uparrow}^\dagger a_{j\downarrow}^\dagger a_{l\downarrow} a_{k\uparrow} \\ &+ \frac{1}{4} \sum_{ijkl} \left(v_{ijkl}^{\alpha\alpha\alpha\alpha} a_{i\uparrow}^\dagger a_{j\uparrow}^\dagger a_{l\uparrow} a_{k\uparrow} + v_{ijkl}^{\beta\beta\beta\beta} a_{i\downarrow}^\dagger a_{j\downarrow}^\dagger a_{l\downarrow} a_{k\downarrow} \right). \end{aligned} \quad (5.5)$$

After applying the Gutzwiller transformation, the Hamiltonian can be written as,

$$\begin{aligned} \overline{H} = & \sum_{ij} \left(t_{ij}^{\alpha\alpha} \rho_{ji}^{(\alpha)} \sigma_{ij}^{(\beta)} + t_{ij}^{\beta\beta} \rho_{ji}^{(\beta)} \sigma_{ij}^{(\alpha)} \right) + \sum_{ijkl} v_{ijkl}^{\alpha\beta\alpha\beta} \Xi_{ijkl}^{(\alpha)} \Xi_{ijkl}^{(\beta)} \\ & + \frac{1}{4} \sum_{ijkl} \left(v_{ijkl}^{\alpha\alpha\alpha\alpha} \Gamma_{ijkl}^{(\alpha)} \Lambda_{ijkl}^{(\beta)} + v_{ijkl}^{\beta\beta\beta\beta} \Gamma_{ijkl}^{(\beta)} \Lambda_{ijkl}^{(\alpha)} \right), \end{aligned} \quad (5.6)$$

where we use the operator definitions:

$$\rho_{ij}^{(\alpha)} = a_{j\uparrow}^\dagger a_{i\uparrow}, \quad (5.7)$$

$$\sigma_{ij}^{(\alpha)} = (1 + \bar{g}_i n_{i\uparrow})(1 + g_j n_{j\uparrow}), \quad (5.8)$$

$$\Gamma_{ijkl}^{(\alpha)} = a_{i\uparrow}^\dagger a_{j\uparrow}^\dagger a_{l\uparrow} a_{k\uparrow}, \quad (5.9)$$

$$\Lambda_{ijkl}^{(\alpha)} = (1 + \bar{g}_i n_{i\uparrow})(1 + \bar{g}_j n_{j\uparrow})(1 + g_k n_{k\uparrow})(1 + g_l n_{l\uparrow}), \quad (5.10)$$

$$\Xi_{ijkl}^{(\alpha)} = (1 + \bar{g}_i n_{i\uparrow}) a_{j\uparrow}^\dagger a_{l\uparrow} (1 + g_k n_{k\uparrow}), \quad (5.11)$$

$$\Xi_{ijkl}^{(\alpha)} = a_{i\uparrow}^\dagger (1 + \bar{g}_j n_{j\uparrow})(1 + g_l n_{l\uparrow}) a_{k\uparrow}. \quad (5.12)$$

We can see that the transformation has the effect of screening the terms in the Hamiltonian with the site-dependant screening coefficients g_i and \bar{g}_i along with the expectation value of an operator composed of terms related to the density of the opposite spin at each site.

The equations for α are constructed as before and simply limited to the same double-occupancy operators as in J ,

$$0 = \langle n_{m\uparrow} n_{m\downarrow} (\overline{H} - E) \rangle. \quad (5.13)$$

It can be seen above that the Hamiltonian will contain terms with up to six-body interactions and when evaluated in 5.13 will result in eight-body terms. The size of

the evaluated terms can be reduced in the case of RHF and UHF wavefunctions as the spin up and down sectors of the wavefunction are separated in the Hilbert space and do not mix. To demonstrate this, if we have some n -body matrix element,

$$\langle \Phi | a_{i_1 \sigma_1}^\dagger a_{i_2 \sigma_2}^\dagger \cdots a_{i_n \sigma_n}^\dagger a_{j_1 \gamma_1} a_{j_2 \gamma_2} \cdots a_{j_n \gamma_n} | \Phi \rangle, \quad (5.14)$$

we can evaluate it in the basis of $|\Phi\rangle$. The form of an RHF or UHF wavefunction does not mix the spin sectors, so we can write,

$$\begin{aligned} a_{i\alpha}^\dagger &= \sum_a A_{io}^* c_{o\alpha}^\dagger + \sum_v A_{iv}^* c_{v\alpha}^\dagger, \\ a_{i\beta}^\dagger &= \sum_a B_{io}^* c_{o\beta}^\dagger + \sum_v B_{iv}^* c_{v\beta}^\dagger, \end{aligned} \quad (5.15)$$

where $c_{o\sigma}^\dagger$ and $c_{v\sigma}$ annihilate the Slater determinant $|\Phi\rangle$. For RHF $A = B$, and for UHF $A \neq B$. In the case of GHF, the transformation cannot be written this way as the Slater determinant basis will mix the up and down spins. As the operators have been written as linear combinations of $\{c, c^\dagger\}$, we can use Wick's theorem to evaluate the matrix where the only non-zero contractions are,

$$\langle c_{o\sigma}^\dagger c_{o'\gamma} \rangle = \delta_{o\sigma'} \delta_{\sigma\gamma}, \quad \langle c_{v\sigma} c_{v'\gamma}^\dagger \rangle = \delta_{vv'} \delta_{\sigma\gamma}. \quad (5.16)$$

Because there is no spin mixing in the transformation between $\{a, a^\dagger\}$ and $\{c, c^\dagger\}$, there are no contractions between different spin sectors when evaluating matrix elements. The only non-zero contractions will be those that share the same spin index. This means that Equation 5.14 can be evaluated as a product of the expectation

values of each spin sector (up to a possible sign due to reordering of the operators),

$$\langle \Phi_\alpha | a_{i_1\alpha}^\dagger a_{i_2\alpha}^\dagger \cdots a_{i_{n_\alpha}\alpha}^\dagger a_{j_1\alpha} a_{j_2\alpha} \cdots a_{j_{n_\alpha}\alpha} | \Phi_\alpha \rangle \langle \Phi_\beta | a_{i_1\beta}^\dagger a_{i_2\beta}^\dagger \cdots a_{i_{n_\beta}\beta}^\dagger a_{j_1\beta} a_{j_2\beta} \cdots a_{j_{n_\beta}\beta} | \Phi_\beta \rangle, \quad (5.17)$$

where n_α and n_β are the number of spin up and down terms respectively, and $|\Phi_\alpha\rangle$ and $|\Phi_\beta\rangle$ simply indicate that we only need consider the spin up and down terms in the wavefunction respectively. The n-body expectation values are now the product of effective n_α and n_β -body expectation values, reducing the number of contractions from $n!$ to $n_\alpha! + n_\beta!$. With this we can see that the largest operator in 5.13 can be evaluated as,

$$\langle \Phi | n_{m\uparrow} n_{m\downarrow} \Gamma_{ijkl}^{(\alpha)} \Lambda_{ijkl}^{(\beta)} | \Phi \rangle = \langle \Phi_\alpha | n_{m\uparrow} \Gamma_{ijkl}^{(\alpha)} | \Phi_\alpha \rangle \langle \Phi_\beta | n_{m\downarrow} \Lambda_{ijkl}^{(\beta)} | \Phi_\beta \rangle. \quad (5.18)$$

This has reduced the evaluation of a term containing an eight-body operator to the evaluation of a five-body and three-body operator.

5.2 Orbital, and Basis Optimization Equations

The Gutzwiller factor on its own is no sufficient for the accurate evaluation of general systems. We need to consider further optimization for the correlation and reference bases in order for the screening to be most effective. We demonstrated earlier that the Gutzwiller correlation factor can produce accurate results for the Hubbard model when the reference Slater determinant is optimized. With this in mind, we can

construct a set of equations for orbital optimization,

$$0 = \frac{\partial}{\partial \kappa_{OV}^*} \frac{\langle \Phi | e^{-K} (1 + \tilde{J}) \bar{H} e^K | \Phi \rangle}{\langle \Phi | e^{-K} (1 + \tilde{J}) e^K | \Phi \rangle} \Big|_{\kappa=0} \quad (5.19)$$

$$0 = \langle \Phi | [(1 + \tilde{J}) \bar{H}, c_V^\dagger c_O - c_O^\dagger c_V] | \Phi \rangle - E \langle \Phi | [\tilde{J}, c_V^\dagger c_O - c_O^\dagger c_V] | \Phi \rangle \quad (5.20)$$

The exact expressions for these conditions are long and uninformative, but with the construction of some intermediates, they can be evaluated with $\mathcal{O}(N^5)$ scaling. However, this on its own is not sufficient to significantly improve the energy and grant the method size consistency. For this, we also need to consider the basis in which the correlation factor is written. In the Hubbard results previously, the on-site basis is a clear choice for the correlation factor. In molecules, this choice is not so obvious. It seems that the atomic orbitals may be a good choice, but there is no reason to assume that this is the best solution. It also cannot be the canonical orbitals of the reference as they are eigenstates of the Jastrow and the correlation factor would simply evaluate to a constant value and introduce no changes to the mean-field calculation. This means that a second basis optimization is required. The condition for the correlation basis can be constructed as stationarity condition for rotations of the correlators ($J' = e^{-K'} J e^{K'}$, $\tilde{J}' = e^{-K'} \tilde{J} e^{K'}$),

$$0 = \frac{\partial}{\partial \kappa'_{IJ}} \frac{\langle \Phi | e^{-K'} (1 + \tilde{J}') e^{-J} e^{K'} H e^{-K'} e^J e^{K'} | \Phi \rangle}{\langle \Phi | e^{-K'} (1 + \tilde{J}') e^{K'} | \Phi \rangle} \Big|_{\kappa'=0} \quad (5.21)$$

If this is satisfied at the same time as the reference optimization, then the rotations next to the reference determinant can be absorbed into the orbital optimization and we only need consider the rotations about the Hamiltonian. this means we need only

consider the condition,

$$0 = \left. \frac{\partial}{\partial \kappa'_{IJ}} \frac{\langle \Phi | (1 + \tilde{J}) e^{-J} e^{K'} H e^{-K'} e^J | \Phi \rangle}{\langle \Phi | (1 + \tilde{J}) | \Phi \rangle} \right|_{\kappa'=0} \quad (5.22)$$

$$0 = \langle \Phi | (1 + \tilde{J}) e^{-J} [H, c_I^\dagger c_J - c_J^\dagger c_I] e^J | \Phi \rangle \quad (5.23)$$

While the orbital optimization condition (Eq. 5.20) only need consider the mixing of the occupied and virtual orbitals, the rotation, K' , for the correlator basis optimization must be a full rotation, not simply particle-hole mixing. While the idea is to rotate the basis of the correlation factor, the problem must be cast in the correlator basis to maintain tractable evaluation. This means that the correlator basis optimization is best run as an optimization of the basis for the Hamiltonian. This means that we change the Hamiltonian via a basis transformation in order to satisfy Eq. 5.23,

$$H[\kappa'] = e^{-K'} H e^{K'}, \quad (5.24a)$$

$$H[\kappa'] = \sum_{PQ} t[\kappa']_{PQ} a_Q^\dagger a_P + \frac{1}{4} \sum_{PQRS} v[\kappa']_{PQRS} a_P^\dagger a_Q^\dagger a_S a_R, \quad (5.24b)$$

$$t[\kappa']_{PQ} = \sum_{IJ} C_{IP}^* t_{IJ} C_{JQ}, \quad (5.24c)$$

$$v[\kappa']_{PQRS} = \sum_{IJKL} C_{IP}^* C_{JQ}^* v_{IJKL} C_{KR} C_{LS}. \quad (5.24d)$$

where $C = e^{K'}$. As long as the transformation C does not mix the spin blocks of the Hamiltonian, we can still separate the expectation values by spin sectors.

If we examine the energy of the hydrogen system in Figure 5.1 when both the correlation basis and the reference are optimized, we can see that there is significant improvement over the results when only the correlator amplitudes and reference orbitals are optimized. The improved correlation operator can now properly separate the two

hydrogen dimers and results in the exact energy for dissociation. The dissociated energy is exact for the separated dimers due to the Gutzwiller transformation being exact for H_2 , just as LAST was exact for two-particle Hubbard lattices in Figure 3.1. The energy at the strongly correlated square configuration is not exact, but now recovers more than 70% of the missing correlation energy.

While this result is promising, there are significant hurdles before this can be used as a general method. Primarily,

these equations are highly unstable. Attempts to obtain benchmark results on other systems have so far led to divergent calculations. The Hubbard calculations in Chapter 4 suffered from similar instabilities when using orbital gradients. Molecular calculation may be solved in a similar way by using a biorthogonal reference, but the number of terms needed to construct the appropriate Fock matrix is not to be taken lightly. While still only scaling as $\mathcal{O}(N^5)$, the largest operator in the energy would result in a derivative with more than 600 terms in the Fock matrix. While this could be improved somewhat by careful grouping of terms in the expansion, this will still result in a much longer and more error prone algebra exercise than the orbital gradients (Eq 5.20). This may however be a good candidate for recent developments in operator algebra software and may yield a simple, effective parameterization of correlated wavefunctions.

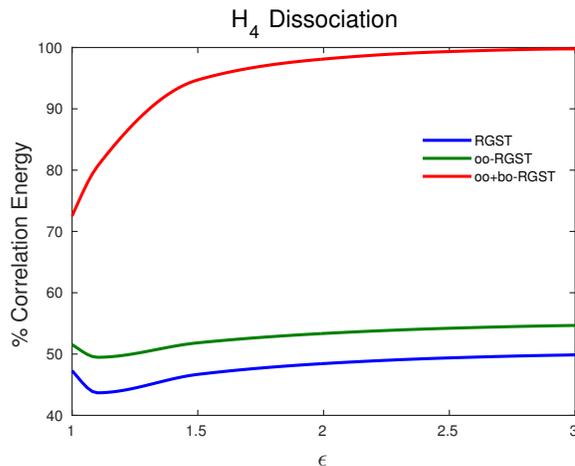


Figure 5.1 : H_4 square with dimensions $1 \times \epsilon$ Angstroms dissociating into two H_2 dimers using the GST in the 3-21G basis including RHF as the reference (RGST), orbital optimization (oo-RGST), and orbital and correlator basis optimization (oo+bo-RGST).

Chapter 6

Conclusions

We have presented similarity transformations generated by exponentials of Hermitian on-site operators resulting in a Hausdorff series which can be resummed and leads to expressions that can be easily evaluated with polynomial cost. Results from this model are in very good agreement with the variational energies, indicating it is a cost-effective way of treating wave functions of the form $e^J|\Phi\rangle$. Results for 1D and 2D systems are of high quality with little computational effort. Our method is size extensive as long as the reference wavefunction is, preserves symmetries that commute with J , and is an alternative to variational Monte Carlo sampling with no stochastic error. The strategy adopted here represents a reasonable approach to optimizing wave functions of the form considered in this work without the need to evaluate the variational energy, which is combinatorial in cost if computed exactly or gains statistical error if calculated via Monte Carlo. Our method could improve our ability to control finite size effects and study impurities directly, particularly in the weakly correlated regime.

By evaluating the transformed Hamiltonian with the more advance projected Hartree-Fock wavefunction, the results are further improved in the strongly correlated regimes with only double occupancy screening. In addition, the calculated double occupancies are consistently close to the best available data, and the errors do not vary greatly for different interaction strengths as in the mean-field calculations. The results are similar to variational Monte Carlo calculations with similar wavefunc-

tions, indicating our method is a good approximation to the variational solution. The projected wavefunction corrects much of the error in the strongly correlated cases, but the additional energy from projection suffers from lack of size extensivity.

The results are comparable to and sometimes better than UCCSD, a much more costly method scaling at $\mathcal{O}(M^5)$ for the Hubbard model versus $\mathcal{O}(M^3)$ in UGST. Some of the current shortfalls could be addressed by evaluating UGST with more advanced wavefunctions and Jastrow factors. Long-range terms can be included in the transformation to improve the description of the correlation functions. Projected GHF and multireference projected wavefunctions are likely candidates that build on the current results as they better address doped and strongly correlated systems on larger lattices [7]. They would also provide a framework to calculate excited states in order to explore the low-lying energy spectrum.

While this is an effective method for the Hubbard model, there is still some work to be done for LAST to be an effective method for general systems. We know that correlation factors of this form can be effective in molecular calculations when calculated variationally [21], but the projective form suffers from stability problems. The biorthogonal ansatz use for UGST may serve to ameliorate this problem, but the calculation of the Fock matrix is a limiting factor. The number of terms in the energy expression when it is fully expanded in the single particle density matrix makes the evaluation of the appropriate derivative unfeasible in a reasonable timeframe. However, with recent developments in symbolic manipulation tools for operator algebras, this derivation could be performed via computer. If this improves the stability of the method, this may become an effective tool for electronic structure calculations, as simple double occupancy screening with both orbital and basis optimization significantly improves the energy of the H_4 benchmark in Chapter 5.

Appendix A

Evaluation of PHF Expectation Values

Here is a demonstration of the evaluation of an expectation value with a SUHF wavefunction. This example can easily be applied to the evaluation of any operator that preserves the spin symmetries. Operators that do not preserve the spin symmetries can also be evaluated but require projection on both the left and right reference determinants and some or all of the spin integrations of the full projection (Eq. 2.31).

We will simply consider the SUHF energy for this example,

$$E = \frac{\langle \Phi | HP | \Phi \rangle}{\langle \Phi | P | \Phi \rangle}, \quad (\text{A.1})$$

where we use the SUHF form of the projected wavefunction,

$$P|\Phi\rangle = \frac{2s+1}{2} \int_0^\pi d\beta \sin \beta d_{mm}^s(\beta) e^{i\beta S_y} |\Phi\rangle. \quad (\text{A.2})$$

To evaluate this, we simply need to integrate the energy over a set of discretized rotation angles (β_i). This is done by first writing the mean-field energy as a function of the density (Eq. 2.10),

$$E[\rho] = \sum_{IJ} t_{IJ} \rho_{JI} + \frac{1}{2} \sum_{IJKL} v_{IJKL} \rho_{KI} \rho_{LJ}. \quad (\text{A.3})$$

Then, at each integration point, we write the transition density between the deformed

determinant $|\Phi\rangle$ and the rotated determinant $e^{i\beta S_y}|\Phi\rangle$.

$$\Lambda(\beta) = \frac{1}{2} \begin{pmatrix} 0 & \beta I \\ -\beta I & 0 \end{pmatrix}, \quad (\text{A.4a})$$

$$C(\beta) = e^{\Lambda(\beta)} C, \quad (\text{A.4b})$$

$$S(\beta) = C^\dagger C(\beta), \quad (\text{A.4c})$$

$$\rho(\beta) = C(\beta) S^{-1}(\beta) C^\dagger, \quad (\text{A.4d})$$

where C is the matrix of coefficients for the occupied orbitals of the deformed determinant (Eq. 2.7). The expectation values now can be written as,

$$\langle \Phi | HP | \Phi \rangle = \frac{2s+1}{2} \int_0^\pi d\beta \sin \beta d_{mm}^s(\beta) \det(S(\beta)) E[\rho(\beta)], \quad (\text{A.5a})$$

$$\langle \Phi | P | \Phi \rangle = \frac{2s+1}{2} \int_0^\pi d\beta \sin \beta d_{mm}^s(\beta) \det(S(\beta)). \quad (\text{A.5b})$$

These integrals are then evaluated with Gaussian quadrature.

Appendix B

Additional Expressions and Results for LAST

In this portion we will present some more detailed expressions that were touched on in Chapters 3 and 4. These will primarily include derived forms of the expressions that were previously left as commutators or derivatives. Not all of the expectation values will be carried through to the full expansion however, as there are several equations where the result is simply too long to be either useful or informative to display. In these cases the expectation value of each term was generated directly into executable code via a recursive algorithm.

B.1 The LAST transformation

Here we show the evaluation of the LAST Transformation in Equation 3.5,

$$e^{-J} a_{k\sigma}^\dagger e^J = e^{-J_{k\sigma}} a_{k\sigma}^\dagger, \quad (\text{B.1})$$

with the definition

$$J_{k\sigma} = \sum_{i\sigma'} \alpha_{k\sigma, i\sigma'} n_{i\sigma'}. \quad (\text{B.2})$$

Expand in the commutator series

$$e^{-J} a_{k\sigma}^\dagger e^J = a_{k\sigma}^\dagger + [a_{k\sigma}^\dagger, J] + \frac{1}{2} [[a_{k\sigma}^\dagger, J], J] + \dots \quad (\text{B.3})$$

The first commutator is

$$\left[a_{k\sigma}^\dagger, J \right] = \frac{1}{2} \sum_{j\sigma'', i\sigma'} \alpha_{j\sigma'', i\sigma'} \left[a_{k\sigma}^\dagger, n_{i\sigma'} n_{j\sigma''} \right] \quad (\text{B.4})$$

$$\begin{aligned} &= - \sum_{i\sigma'} \alpha_{k\sigma, i\sigma'} n_{i\sigma'} a_{k\sigma}^\dagger \\ &= -J_{k\sigma} c_{k\sigma}^\dagger \end{aligned} \quad (\text{B.5})$$

The second commutator is

$$\begin{aligned} \left[\left[a_{p\sigma}^\dagger, J \right], J \right] &= -J_{k\sigma} \left[a_{k\sigma}^\dagger, J \right] \\ &= J_{k\sigma}^2 a_{k\sigma}^\dagger \end{aligned} \quad (\text{B.6})$$

because

$$\left[J_{k\sigma}, J \right] = 0. \quad (\text{B.7})$$

From here, the main result follows immediately as further commutators will simply result in further factors of $-J_{K\sigma}$.

B.2 LAST Expectation Values

The method for evaluating the expectation value for LAST is already outlined in Eq. 3.16, but we include some explicit expectation values here for demonstration.

Once again, the energy expression for LAST on the Hubbard Hamiltonian is,

$$E_{LAST} = -t \sum_{\langle ij \rangle, \sigma} \langle \Phi | e^{-J_{i\sigma}} a_{i\sigma}^\dagger a_{j\sigma} e^{J_{j\sigma}} | \Phi \rangle + U \sum_i \langle \Phi | n_{i\uparrow} n_{i\downarrow} | \Phi \rangle \quad (\text{B.8})$$

The interaction term commutes with the correlation factor, so this can be evaluated in mean-field as a expansion in the single-particle density,

$$\langle \Phi | n_{i\uparrow} n_{i\downarrow} | \Phi \rangle = \rho_{i\uparrow, i\uparrow} \rho_{i\downarrow, i\downarrow} - \rho_{i\uparrow, i\downarrow} \rho_{i\downarrow, i\uparrow}. \quad (\text{B.9})$$

The hopping term, as mentioned in Chapter 3, requires the construction of a site-dependant density to incorporate the local modifications brought by the correlation factors. If we write the site dependent density as,

$$\det(S^{(i\sigma, j\sigma')}) \rho_{k\gamma, l\gamma'}^{(i\sigma, j\sigma')} = \langle \Phi | e^{-J_{i\sigma}} a_{l\gamma'}^\dagger a_{k\gamma} e^{J_{j\sigma'}} | \Phi \rangle \quad (\text{B.10})$$

with the corresponding overlap matrix $S^{(i\sigma, j\sigma')}$, the LAST energy expression becomes,

$$E_{LAST} = -t \sum_{\langle ij \rangle, \sigma} \det(S^{(i\sigma, j\sigma)}) \rho_{i\sigma, j\sigma}^{(i\sigma, j\sigma)} + U \sum_i (\rho_{i\uparrow, i\uparrow} \rho_{i\downarrow, i\downarrow} - \rho_{i\uparrow, i\downarrow} \rho_{i\downarrow, i\uparrow}). \quad (\text{B.11})$$

The amplitude equation can be constructed in the same manner, The main difference being that we now need to perform an expansion with Wick's theorem in terms of the local density as well,

$$\begin{aligned} R_{i\sigma, j\sigma'} = & -t \sum_{\langle kl \rangle, \gamma} \langle \Phi | n_{i\sigma} n_{j\sigma'} e^{-J_{k\gamma}} c_{k\gamma}^\dagger c_{l\gamma} e^{J_{l\gamma}} | \Phi \rangle + U \sum_k \langle \Phi | n_{i\sigma} n_{j\sigma'} n_{k\uparrow} n_{k\downarrow} | \Phi \rangle \\ & - E_{LAST} \langle \Phi | n_{i\sigma} n_{j\sigma'} | \Phi \rangle. \end{aligned} \quad (\text{B.12})$$

The interaction is again a simple, albeit long, expansion in the mean-field density, and the hopping term is written in terms of the local density,

$$\begin{aligned}
\langle \Phi | n_{i\sigma} n_{j\sigma'} e^{-J_{k\gamma} c_{k\gamma}^\dagger c_{l\gamma}} e^{J_{l\gamma}} | \Phi \rangle &= \det(S^{(k\gamma, l\gamma)}) (\tag{B.13} \\
&\rho_{i\sigma, i\sigma}^{(k\gamma, l\gamma)} (\rho_{j\sigma', j\sigma'}^{(k\gamma, l\gamma)} \rho_{l\gamma, k\gamma}^{(k\gamma, l\gamma)} - \rho_{l\gamma, j\sigma'}^{(k\gamma, l\gamma)} \rho_{j\sigma', k\gamma}^{(k\gamma, l\gamma)}) \\
&+ \rho_{j\sigma', i\sigma}^{(k\gamma, l\gamma)} (\rho_{i\sigma, k\gamma}^{(k\gamma, l\gamma)} \rho_{l\gamma, j\sigma'}^{(k\gamma, l\gamma)} - \rho_{i\sigma, j\sigma'}^{(k\gamma, l\gamma)} \rho_{l\gamma, k\gamma}^{(k\gamma, l\gamma)}) \\
&+ \rho_{l\gamma, i\sigma}^{(k\gamma, l\gamma)} (\rho_{i\sigma, j\sigma'}^{(k\gamma, l\gamma)} \rho_{j\sigma', k\gamma}^{(k\gamma, l\gamma)} - \rho_{j\sigma', j\sigma'}^{(k\gamma, l\gamma)} \rho_{i\sigma, k\gamma}^{(k\gamma, l\gamma)}) \\
&+ \rho_{i\sigma, i\sigma}^{(k\gamma, l\gamma)} \rho_{l\gamma, j\sigma'}^{(k\gamma, l\gamma)} \delta_{j\sigma', k\gamma} + \rho_{l\gamma, i\sigma}^{(k\gamma, l\gamma)} \delta_{i\sigma, j\sigma'} \delta_{j\sigma', k\gamma} \\
&+ \rho_{j\sigma', i\sigma}^{(k\gamma, l\gamma)} (\rho_{l\gamma, k\gamma}^{(k\gamma, l\gamma)} \delta_{i\sigma, j\sigma'} - \rho_{l\gamma, j\sigma'}^{(k\gamma, l\gamma)} \delta_{i\sigma, k\gamma}) \\
&- \rho_{l\gamma, i\sigma}^{(k\gamma, l\gamma)} (\rho_{j\sigma', k\gamma}^{(k\gamma, l\gamma)} \delta_{i\sigma, j\sigma'} + \rho_{i\sigma, j\sigma'}^{(k\gamma, l\gamma)} \delta_{j\sigma', k\gamma}) \\
&+ \rho_{l\gamma, i\sigma}^{(k\gamma, l\gamma)} \rho_{j\sigma', j\sigma'}^{(k\gamma, l\gamma)} \delta_{i\sigma, k\gamma}).
\end{aligned}$$

B.3 GST Expectation Values

When the calculation is limited to the Gutzwiller transformation, we no longer need to construct the transition density for each site. Instead we can write out the three-body Hamiltonian. This increases the size of the operators, and therefore the number of terms in the expansion, but serves to reduce the scaling of the calculation as we no longer need to construct a new overlap at each site. For UGST where we need not consider the mixing of the spin up and down electrons, The expectation values for the Hubbard model are relatively simple. Here the spin up and down sectors of the

full density matrix are written as ρ_{ij}^α and ρ_{ij}^β respectively.

$$\begin{aligned} \langle \Phi' | \bar{H} | \Phi \rangle = & -t \sum_{\langle ij \rangle} \left(\left(1 + \bar{g}_i \rho_{ii}^\beta + g_j \rho_{jj}^\beta + \bar{g}_i g_j (\rho_{ii}^\beta \rho_{jj}^\beta - \rho_{ji}^\beta \rho_{ij}^\beta) \right) \rho_{ji}^\alpha \right. \\ & \left. + (\alpha \leftrightarrow \beta) \right) + U \sum_i \rho_{ii}^\alpha \rho_{ii}^\beta \end{aligned} \quad (\text{B.14})$$

$$\begin{aligned} \langle \Phi' | n_{K\uparrow} n_{K\downarrow} (\bar{H} - E) | \Phi \rangle = & -t \sum_{\langle ij \rangle} \left(\left(\rho_{kk}^\beta + \bar{g}_i \langle \Phi' | n_{k\downarrow} n_{i\downarrow} | \Phi \rangle + g_j \langle \Phi' | n_{k\downarrow} n_{j\downarrow} | \Phi \rangle \right. \right. \\ & \left. \left. + \bar{g}_i g_j \langle \Phi' | n_{k\downarrow} n_{i\downarrow} n_{j\downarrow} | \Phi \rangle \right) (\delta_{ik} \rho_{ki}^\alpha + \rho_{kk}^\alpha \rho_{ji}^\alpha - \rho_{jk}^\alpha \rho_{ki}^\alpha) \right. \\ & \left. + (\alpha \leftrightarrow \beta) \right) + U \sum_i \langle \Phi' | n_{k\uparrow} n_{i\uparrow} | \Phi \rangle \langle \Phi' | n_{k\downarrow} n_{i\downarrow} | \Phi \rangle \\ & - \rho_{kk}^\alpha \rho_{kk}^\beta E \end{aligned} \quad (\text{B.15})$$

$$\langle \Phi' | n_{i\downarrow} n_{j\downarrow} | \Phi \rangle = \delta_{ij} \rho_{ji}^\beta + \rho_{ii}^\beta \rho_{jj}^\beta - \rho_{ji}^\beta \rho_{ij}^\beta \quad (\text{B.16})$$

$$\begin{aligned} \langle \Phi' | n_{k\downarrow} n_{i\downarrow} n_{j\downarrow} | \Phi \rangle = & \rho_{kk}^\beta (\delta_{ij} \rho_{ji}^\beta + \rho_{ii}^\beta \rho_{jj}^\beta - \rho_{ji}^\beta \rho_{ij}^\beta) \\ & + \rho_{ik}^\beta (\delta_{ik} \rho_{jj}^\beta - \delta_{jk} \rho_{ji}^\beta - \rho_{jj}^\beta \rho_{ki}^\beta + \rho_{ji}^\beta \rho_{kj}^\beta) \\ & + \rho_{jk}^\beta (\delta_{ij} \delta_{ik} - \delta_{ij} \rho_{ki}^\beta - \delta_{ik} \rho_{ij}^\beta + \delta_{jk} \rho_{ii}^\beta + \rho_{ij}^\beta \rho_{ki}^\beta - \rho_{ii}^\beta \rho_{kj}^\beta) \end{aligned} \quad (\text{B.17})$$

B.4 Additional Results

The following tables contain calculated correlation energies, double occupancies, and reference data with coupled-cluster singles and doubles (CCSD), exact diagonalization (ED), auxiliary-field quantum Monte Carlo (AFQMC), density matrix embedding theory (DMET), density matrix renormalization group theory (DMRG), and diffusion Monte Carlo based on a fixed-node approximation (FN). Uncertainties for the last digit from the methods or thermodynamic limit extrapolations are given in parentheses.

N_o	U	E_{RHF}	E_{UHF}	E_{RJ}	E_{UJ}	E_{ED}	$\%E_c(RJ)$	$\%E_c(UJ)$
4	4	-0.6875	-0.7053	-0.7120	-0.7194	-0.7206	73.93	96.33
4	8	-0.6250	-0.6663	-0.6962	-0.7048	-0.7076	86.19	96.55
4	12	-0.5625	-0.6345	-0.6878	-0.6959	-0.7003	90.94	96.84
6	4	-0.8594	-0.9000	-0.9105	-0.9274	-0.9460	58.97	78.55
6	8	-0.7187	-0.8240	-0.8723	-0.8927	-0.9202	76.24	86.36
6	12	-0.5781	-0.7730	-0.8526	-0.8733	-0.9061	83.69	89.99
8	4	-1.0000	-1.0249	-1.0811	-1.0885	-1.0959	84.55	92.24
8	8	-0.7500	-0.8469	-1.0052	-1.0031	-1.0288	91.52	90.77
8	12	-0.5000	-0.7572	-0.9767	-0.9347	-0.9941	96.47	87.97
10	4	-1.1094	-1.1094	-1.2162	-1.2162	-1.2238	93.35	93.35
10	8	-0.7187	-0.8490	-1.0678	-1.0157	-1.0944	92.93	79.06
10	12	-0.3281	-0.7393	-0.9787	-0.9284	-1.0284	92.91	85.72
12	4	-0.9375	-0.9861	-1.0837	-1.0898	-1.1080	85.75	89.30
12	8	-0.3750	-0.7091	-0.8694	-0.8483	-0.9328	88.63	84.85
12	12	0.1875	-0.5967	-0.8508	-0.7491	-0.8512	99.96	90.17
14	4	-0.7344	-0.8808	-0.9018	-0.9595	-0.9840	67.06	90.18
14	8	0.0313	-0.5921	-0.5354	-0.6691	-0.7418	73.30	90.60
14	12	0.7969	-0.4744	-0.3158	-0.5488	-0.6282	78.08	94.43
16	4	-0.5000	-0.7854	-0.6931	-0.8270	-0.8514	54.94	93.06
16	8	0.5000	-0.4619	-0.2235	-0.4873	-0.5293	70.29	95.92
16	12	1.5000	-0.3208	-0.0447	-0.3326	-0.3745	82.41	97.76

Table B.1 : Correlated energies for 4x4 Hubbard lattices on RHF (E_{RJ}) and UHF (E_{UJ}) references and the portion of the correlation energy recovered (E_c), exact energies (E_{ED}) taken from [36]. E_{UJ} includes spin-density correlators ($S_i^z N_j$).

N_o	U	E_{UHF}	Δ_{SUHF}	Δ_{UGST}	Δ_{SUGST}	Δ_{UCCSD}	Δ_{ED}
14	1	-1.3277	-0.0015	-0.0074	-0.0084	-0.0085	-0.0093
14	2	-1.1644	-0.0069	-0.0271	-0.0309	-0.0309	-0.0338
14	3	-1.0138	-0.0168	-0.0540	-0.0615	-0.0606	-0.0674
14	4	-0.8808	-0.0175	-0.0813	-0.0930	-0.0907	-0.1032
14	6	-0.6988	-0.0382	-0.0979	-0.1173	-0.1083	-0.1400
14	8	-0.5921	-0.0352	-0.0877	-0.1121	-0.1056	-0.1497
14	10	-0.5227	-0.0315	-0.0799	-0.1019	-0.1022	-0.1527
14	12	-0.4819	-0.0256	-0.0689	-0.0880	-0.0869	-0.1463
16	1	-1.2909	-0.0022	-0.0066	-0.0082	-0.0077	-0.0087
16	2	-1.0973	-0.0098	-0.0224	-0.0281	-0.0252	-0.0288
16	3	-0.9267	-0.0220	-0.0393	-0.0505	-0.0435	-0.0506
16	4	-0.7854	-0.0338	-0.0494	-0.0667	-0.0552	-0.0660
16	6	-0.5862	-0.0426	-0.0475	-0.0736	-0.0590	-0.0733
16	8	-0.4619	-0.0410	-0.0354	-0.0645	-0.0536	-0.0674
16	10	-0.3792	-0.0373	-0.0250	-0.0536	-0.0477	-0.0601
16	12	-0.3208	-0.0336	-0.0178	-0.0449	-0.0426	-0.0537

Table B.2 : Correlation energies per site with respect to UHF for 4×4 Hubbard lattices with spin quantum numbers $s = m = 0$ and N_o electrons. Exact energies (ED) taken from [38].

n	U	E_{UHF}	Δ_{UGST}	Δ_{SUGST}	$\Delta_{UCCSD}^{a,b}$	Δ_{TDL}^a	
0.8	2	-1.2678	-0.0358	-0.0361	-0.0416	AFQMC	-0.038(2)
						DMET	-0.0384(4)
						FN	-0.0366(7)
0.8	4	-0.9774	-0.1061	-0.1087	-0.1151	AFQMC	-0.133(3)
						DMET	-0.131(2)
						DMRG	-0.127(1)
						FN	-0.1258(7)
0.8	6	-0.7933	-0.1335	-0.1373	-0.1275	DMET	-0.184(4)
						FN	-0.174(1)
0.8	8	-0.6815	-0.1386	-0.1424	-0.1431	DMET	-0.20(3)
						FN	-0.196(1)
1.0	2	-1.1354	-0.0286	-0.0302	-0.0318	AFQMC	-0.0409(2)
						DMET	-0.0410(3)
						DMRG	-0.041(1)
						FN	-0.040(1)
1.0	4	-0.7970	-0.0532	-0.0572	-0.0582	AFQMC	-0.0633(2)
						DMET	-0.0634(3)
						DMRG	-0.0635(5)
						FN	-0.0605(3)
1.0	6	-0.5927	-0.0487	-0.0537	-0.0583	AFQMC	-0.0641(3)
						DMET	-0.0635(5)
						DMRG	-0.0638(1)
						FN	-0.0624(1)
1.0	8	-0.4659	-0.0366	-0.0417	-0.0532	AFQMC	-0.0588(2)
						DMET	-0.058(1)
						DMRG	-0.0582(1)
						FN	-0.0573
1.0	12	-0.3225	-0.0181	-0.0229	-0.0422	AFQMC	-0.0468(2)
						DMET	-0.046(1)
						DMRG	-0.0464(1)
						FN	-0.0459

^a Results taken from [39]. ^b Results taken from [40].

Table B.3 : Correlation energies per site with respect to UHF for 10×10 lattices with average occupancy n .

n	U	E_{UHF}	Δ_{SUGST}	Δ_{UCCSD}^a	Δ_{TDL}^a	
0.8	2	-1.2602	-0.0376	-0.0463 ^b	AFQMC	-0.046(2)
					DMET	-0.0460(4)
					FN	-0.0442(7)
0.8	4	-0.9762	-0.1026	-0.1106 ^b	AFQMC	-0.134(3)
					DMET	-0.132(2)
					DMRG	-0.128(1)
					FN	-0.1270(7)
0.8	6	-0.7843	-0.1416	-0.1457 ^b	DMET	-0.193(4)
					FN	-0.183(1)
0.8	8	-0.6772	-0.1432	-0.1461 ^b	DMET	-0.20(3)
					FN	-0.200(1)
1.0	2	-1.1389	-0.0310	-0.0364(4)	AFQMC	-0.0374(2)
					DMET	-0.0375(3)
					DMRG	-0.037(1)
					FN	-0.036(1)
1.0	4	-0.7978	-0.0530	-0.0568 ^c	AFQMC	-0.0625(2)
					DMET	-0.0626(3)
					DMRG	-0.0627(5)
					FN	-0.0597(3)
1.0	6	-0.5927	-0.0492	-0.0585 ^d	AFQMC	-0.0641(3)
					DMET	-0.0635(5)
					DMRG	-0.0638(1)
					FN	-0.0624(1)
1.0	8	-0.4659	-0.0371	-0.0532 ^d	AFQMC	-0.0588(2)
					DMET	-0.058(1)
					DMRG	-0.0582(1)
					FN	-0.0573
1.0	12	-0.3225	-0.0187	-0.0422 ^d	AFQMC	-0.0468(2)
					DMET	-0.046(1)
					DMRG	-0.0464(1)
					FN	-0.0459

^a Results taken from [39]. ^b 10×16 lattice, ^c 14×14 lattice, ^d 10×10 lattice

Table B.4 : SUGST correlation energies per site for 30×30 square lattices with average occupancy n .

U	UHF	SUGST	UCCSD ^a	TDL ^a	
2	0.2146	0.1943	0.194(2)	AFQMC	0.1923(3)
				DMET	0.1913(4)
				DMRG	0.188(1)
				FN	0.198(1)
4	0.1307	0.1282	0.1268 ^b	AFQMC	0.1261(2)
				DMET	0.1261(1)
				DMRG	0.126(1)
				FN	0.125(1)
6	0.0789	0.0843	0.0807 ^c	AFQMC	0.0810(1)
				DMET	0.0810
				DMRG	0.0809(3)
				FN	0.0803(2)
8	0.0507	0.0568	0.0537 ^c	AFQMC	0.0540(1)
				DMET	0.0540
				DMRG	0.0539(1)
				FN	0.0535(1)
12	0.0252	0.0283	0.0267 ^c	AFQMC	0.0278(1)
				DMET	0.0278
				DMRG	0.0278(1)
				FN	0.0278(2)

^a Results taken from [39]. ^b 12×12 lattice, ^c 10×10 lattice

Table B.5 : Average SUGST double occupancies per site for 30×30 square lattices at half-filling.

Bibliography

- [1] R. J. Bartlett and M. Musiał, “Coupled-cluster theory in quantum chemistry,” *Rev. Mod. Phys.*, vol. 79, pp. 291–352, Feb 2007.
- [2] C. A. Jimnez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, “Projected hartree-fock theory,” *The Journal of Chemical Physics*, vol. 136, no. 16, p. 164109, 2012.
- [3] T. Stein, T. M. Henderson, and G. E. Scuseria, “Seniority zero pair coupled cluster doubles theory,” *The Journal of Chemical Physics*, vol. 140, no. 21, p. 214113, 2014.
- [4] P. Ring and P. Schuck, *The Nuclear Many-Body Problem*. Berlin: Springer, 1980.
- [5] J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems*. Cambridge, MA: MIT, 1985.
- [6] D. Thouless, “Stability conditions and nuclear rotations in the hartree-fock theory,” *Nuclear Physics*, vol. 21, pp. 225 – 232, 1960.
- [7] R. Rodríguez-Guzmán, K. W. Schmid, C. A. Jiménez-Hoyos, and G. E. Scuseria, “Symmetry-projected variational approach for ground and excited states of the two-dimensional hubbard model,” *Phys. Rev. B*, vol. 85, p. 245130, Jun 2012.
- [8] C. A. Jimnez-Hoyos, T. M. Henderson, and G. E. Scuseria, “Generalized hartreefock description of molecular dissociation,” *Journal of Chemical Theory*

- and Computation*, vol. 7, no. 9, pp. 2667–2674, 2011. PMID: 26605457.
- [9] S. F. Boys and N. C. Handy, “The determination of energies and wavefunctions with full electronic correlation,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 310, no. 1500, pp. 43–61, 1969.
- [10] S. D. Głazek and K. G. Wilson, “Renormalization of hamiltonians,” *Phys. Rev. D*, vol. 48, pp. 5863–5872, Dec 1993.
- [11] F. Wegner, “Flow-equations for hamiltonians,” *Annalen der Physik*, vol. 506, no. 2, pp. 77–91, 1994.
- [12] M. Nooijen, “Manybody similarity transformations generated by normal ordered exponential excitation operators,” *The Journal of Chemical Physics*, vol. 104, no. 7, pp. 2638–2651, 1996.
- [13] S. R. White, “Numerical canonical transformation approach to quantum many-body problems,” *The Journal of Chemical Physics*, vol. 117, no. 16, pp. 7472–7482, 2002.
- [14] T. Yanai and G. K.-L. Chan, “Canonical transformation theory for multireference problems,” *The Journal of Chemical Physics*, vol. 124, no. 19, p. 194106, 2006.
- [15] M. Dzierzawa, D. Baeriswyl, and L. M. Martelo, “Variational wave functions for the mott-hubbard transition,” *Helvetica Physica Acta*, vol. 70, pp. 124–140, 1997.
- [16] M. Casula and S. Sorella, “Geminal wave functions with jastrow correlation: A first application to atoms,” *The Journal of Chemical Physics*, vol. 119, no. 13,

- pp. 6500–6511, 2003.
- [17] G. Stollhoff and P. Fulde, “On the computation of electronic correlation energies within the local approach,” *The Journal of Chemical Physics*, vol. 73, no. 9, pp. 4548–4561, 1980.
- [18] B. Edegger, V. N. Muthukumar, and C. Gros, “Gutzwiller-*r*vb theory of high-temperature superconductivity: Results from renormalized mean-field theory and variational monte carlo calculations,” *Advances in Physics*, vol. 56, no. 6, pp. 927–1033, 2007.
- [19] D. Baeriswyl, D. Eichenberger, and M. Menteshashvili, “Variational ground states of the two-dimensional hubbard model,” *New Journal of Physics*, vol. 11, no. 7, p. 075010, 2009.
- [20] F. Gebhard and M. Gutzwiller, “Gutzwiller wave function,” *Scholarpedia*, vol. 4, no. 4, p. 7288, 2009.
- [21] E. Neuscamman, “Communication: A jastrow factor coupled cluster theory for weak and strong electron correlation,” *The Journal of Chemical Physics*, vol. 139, no. 18, p. 181101, 2013.
- [22] T. M. Henderson and G. E. Scuseria, “Linearized jastrow-style fluctuations on spin-projected hartree-fock,” *The Journal of Chemical Physics*, vol. 139, no. 23, p. 234113, 2013.
- [23] D. Tahara and M. Imada, “Variational monte carlo method combined with quantum-number projection and multi-variable optimization,” *Journal of the Physical Society of Japan*, vol. 77, no. 11, p. 114701, 2008.

- [24] S. Tsuneyuki, “Transcorrelated method: Another possible way towards electronic structure calculation of solids,” *Progress of Theoretical Physics Supplement*, vol. 176, p. 134, 2008.
- [25] N. Handy, “On the minimization of the variance of the transcorrelated hamiltonian,” *Molecular Physics*, vol. 21, no. 5, pp. 817–828, 1971.
- [26] S. Ten-no, “A feasible transcorrelated method for treating electronic cusps using a frozen gaussian geminal,” *Chemical Physics Letters*, vol. 330, no. 12, pp. 169 – 174, 2000.
- [27] N. Umezawa and S. Tsuneyuki, “Transcorrelated method for electronic systems coupled with variational monte carlo calculation,” *The Journal of Chemical Physics*, vol. 119, no. 19, pp. 10015–10031, 2003.
- [28] E. Neuscamman, H. Changlani, J. Kinder, and G. K.-L. Chan, “Nonstochastic algorithms for jastrow-slater and correlator product state wave functions,” *Phys. Rev. B*, vol. 84, p. 205132, Nov 2011.
- [29] M. Nooijen, “Can the eigenstates of a many-body hamiltonian be represented exactly using a general two-body cluster expansion?,” *Phys. Rev. Lett.*, vol. 84, pp. 2108–2111, Mar 2000.
- [30] W. Kutzelnigg and D. Mukherjee, “Minimal parametrization of an n -electron state,” *Phys. Rev. A*, vol. 71, p. 022502, Feb 2005.
- [31] P.-D. Fan and P. Piecuch, “The usefulness of exponential wave function expansions employing one- and two-body cluster operators in electronic structure theory: The extended and generalized coupled-cluster methods,” *Advances in Quantum Chemistry*, vol. 51, pp. 1 – 57, 2006.

- [32] R. Gilmore., *Lie Groups, Lie Algebras, and Some of Their Applications*. New York: Dover, 2005.
- [33] A. Anderson, “Canonical transformations in quantum mechanics,” *Annals of Physics*, vol. 232, no. 2, pp. 292 – 331, 1994.
- [34] B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler, A. Hehn, R. Igarashi, S. V. Isakov, D. Koop, P. N. Ma, P. Mates, H. Matsuo, O. Parcollet, G. Pawowski, J. D. Picon, L. Pollet, E. Santos, V. W. Scarola, U. Schollwck, C. Silva, B. Surer, S. Todo, S. Trebst, M. Troyer, M. L. Wall, P. Werner, and S. Wessel, “The alps project release 2.0: open source software for strongly correlated systems,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 05, p. P05001, 2011.
- [35] A. Albuquerque, F. Alet, P. Corboz, P. Dayal, A. Feiguin, S. Fuchs, L. Gamper, E. Gull, S. Grtler, A. Honecker, R. Igarashi, M. Krner, A. Kozhevnikov, A. Luchli, S. Manmana, M. Matsumoto, I. McCulloch, F. Michel, R. Noack, G. Pawowski, L. Pollet, T. Pruschke, U. Schollwck, S. Todo, S. Trebst, M. Troyer, P. Werner, and S. Wessel, “The {ALPS} project release 1.3: Open-source software for strongly correlated systems,” *Journal of Magnetism and Magnetic Materials*, vol. 310, no. 2, Part 2, pp. 1187 – 1193, 2007. Proceedings of the 17th International Conference on MagnetismThe International Conference on Magnetism.
- [36] H. Shi and S. Zhang, “Symmetry in auxiliary-field quantum monte carlo calculations,” *Phys. Rev. B*, vol. 88, p. 125132, Sep 2013.
- [37] H. Shi, C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, G. E. Scuseria, and

- S. Zhang, “Symmetry-projected wave functions in quantum monte carlo calculations,” *Phys. Rev. B*, vol. 89, p. 125129, Mar 2014.
- [38] G. Fano, F. Ortolani, and A. Parola, “Hole-hole effective interaction in the two-dimensional hubbard model,” *Phys. Rev. B*, vol. 42, pp. 6877–6880, Oct 1990.
- [39] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof’ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull, “Solutions of the two-dimensional hubbard model: Benchmarks and results from a wide range of numerical algorithms,” *Phys. Rev. X*, vol. 5, p. 041041, Dec 2015.
- [40] C. A. Jiménez-Hoyos and G. E. Scuseria, “Cluster-based mean-field and perturbative description of strongly correlated fermion systems: Application to the one- and two-dimensional hubbard model,” *Phys. Rev. B*, vol. 92, p. 085101, Aug 2015.
- [41] H. Yokoyama and H. Shiba, “Variational monte-carlo studies of hubbard model. ii,” *Journal of the Physical Society of Japan*, vol. 56, no. 10, pp. 3582–3592, 1987.