Algorithms for the Nuclear Many-Body Problem and Beyond

by

Ashe Hicks

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics - Doctor of Philosophy
Computational Mathematics, Science, and Engineering - Dual Major

2024

Abstract

The nuclear many body problem allows us to take our fundamental understanding of the most basic building blocks of the universe and from them build an understanding of larger and more complicated systems. It is the essential problem of how individual particles form atoms and larger structures. Its applications are varied, and many tools have been developed to address this problem. Despite the breakneck pace of computational development, the nuclear many-body problem still stretches our computational and numerical methods to and beyond their breaking points. In this work, we introduce two algorithms which can help in solving the nuclear many-body problem. First, we introduce trimmed sampling. This is an algorithm which can be used to treat noisy data obtained from highly sensitive calculations, particularly the generalized eigenvalue problem which emerges from a number of techniques. We solve a number of example models for which small errors such as rounding error or statistical noise are sufficient to entirely destroy any usable results, but see that trimmed sampling is able to recover good results from these methods. It does so using Bayesian inference, by applying physics-informed criteria and statistical sampling methods we are able to eliminate any solutions which are non-physical, leaving a more accurate, physically meaningful result. We show ways that this algorithm can be further expanded and enhanced, improving sampling statistics, convergence rate, and accuracy, before demonstrating its performance on the Lipkin model. In the next section, we describe the Projected Cooling algorithm. This is a method whereby we use an analogue of evaporative cooling to calculate the ground state of a system. We show results of projected cooling for several models. Together, this work provides a description of useful algorithms which can be applied to the nuclear many-body problem.

Dedicated to Lore, the greatest adventure buddy I could ever ask for.

Acknowledgements

Lore, my spouse and partner, you have always stood by me, no matter what unfolds around us. When I needed to vent, you listened. When I needed comfort, you comforted me. When I needed an honest kick, you delivered it as gently as you could. You supported me though depression and anxiety, and you build a life with me where I feel safe and secure. You accepted me without hesitation. I am truly lucky to have you in my life. Thank you.

To my friends, Ava, Myles, Jack, and Lamont, thank you for always being available to talk or just relax when I needed to. You ensured that through this whole process I never felt alone or in want of community. You made this time in my life not just survivable, but a time that I look back on fondly. You helped me grow as a person, more than I ever imagined. I wouldn't be the same person I am today without you all around me.

To my parents, thank you for everything you've done for me. You gave me the space and freedom I needed as a child to learn and grow, to explore my own interests, and to become my own person. You helped me not feel pressured to conform to a role which didn't suit me, and started me on the road to eventually figuring out who I am. Your guidance and advice have helped shape me, and your support for me means the world.

To my cats, Salem, Cremini, and Thalia, a begrudging thank you. You have never ceased to distract me from work to remind me of my true purpose: ensuring that your food bowl stays full. Despite my cruel negligence in only feeding you twice per day, you always gave me confort when I was down, and you helped me recharge after a rough day.

Finally, I want to thank my advisor Dean Lee for his incredible guidance and mentorship. He gave me the freedom I needed to explore, but was always there to coach me through the countless dead ends I found myself in and mistakes that I made. His patience, kindness, and constant encouragement are what made this thesis possible. When I was sure that I would never graduate, Dean gave me the encouragement I needed to keep going. And when I finally found success, Dean was there too to help me move forward. In the last months of this thesis, when my personal life was in chaos, Dean showed the same

kindness, encouragement, and acceptance to me personally as he always has academically. I could not imagine a better advisor, and I don't think I could have made it through without him.

Contents

List of Figures

1

Introduction

The nuclear many-body problem describes the emergent phenomena which occur when considering the behavior of a large group of particles, themselves each governed by fundamental laws and interactions. It provides a bridge between the world of fundamental, first prinicples calculations and large scale statistical models which can be measured and interacted with. The many-body problem is a well developed field with many methods which stretch classical computation to its limits. In this work, we present a description of two algorithms which can prove beneficial to the many-body problem, with potential application in other fields.

In chapter 2, we provide the background theory on which much of this work is based. We begin with an overview of quantum mechanics and many-body theory, and proceed to describe how these mechanics can be applied to classical computer systems the limitations in doing so. We will provide an overview of reduced basis methods and similar computational techniques, such as eigenvector continuation and euclidean time projection.

In chapter 3, we describe the bulk of this work, trimmed sampling. Trimmed sampling is an algorithm which may be applied to many existing reduced basis methods to mitigate some of the limitations of those techniques and improve the accuracy of the calculation. It does this by first considering the prior distribution of the parameters which describe that system–in the case of the reduced basis methods we present in this chapter, this is the matrix elements of the reduce basis generalized eigenvalue problem. We then consider a set of physics-informed likelihood criteria and use these likelihood functions to weight samples drawn from the prior distribution. By analyzing the resulting posterior distribution, we gain insight into the system and its range of physically-consistent observables, and in this way we can substantially reduce error in otherwise unsalvageable calculations. We show many examples, beginning with nearly tivial calculations before progressing to some more complicated toy problems, such as the lipkin model and the bose-hubbard model. Along the way, we show several advancements that may be used alongside the basic trimmed sampling algorithm to further refine its capabilities.

In chapter 4, we describe the projected cooling algorithm, a quantum algorithm which can be applied to find the ground state in a quantum system as an analogue of evaporative cooling. It makes use of the fact that unbound states have a tendency to propagate away from their initial region, while low-lying states will remain within a compact region. By driving the excited quantum states to leave, then projecting back onto the local region, we are left with a state with very high overlap with the ground state. We present several models and examples of this algorithm in use.

Finally, in chapter 5 we present our conclusions and perspective on possible future work involving these algorithms to many-body physics and other problems.

2

Background

In this chapter, we give an overview of the methods and fundamentals required for this work. Those familiar with the subjects are encouraged to skip those sections or skim them for a refresher. We will begin with a description of the basics of quantum mechanics, from its initial postulates. We will then describe basic numerical techniques for quantum simulations. We discuss several reduced basis methods and give an overview of monte-carlo methods. Finally, we give a brief introduction to the field of quantum many-body physics.

2.1 Quantum Mechanics

The field of quantum mechanics is a relatively recent development when compared to other fields of physics, such as classical mechanics and electromagnetism. Unlike these more classical fields of study, quantum mechanical observations are probabilistic. Quantum mechancal calculations produce the wavefunction, from which the probability of any given observation may be extracted. One of its earliest formulations was put forward by Erwin Schrödinger in 1926 (1). We will begin with the basic postulates of quantum mechanics, from which the rest of the theory is developed.

2.1.1 Postulates of Quantum Mechanics

The following postulates form the foundation on which the rest of the theory is built. They are as follows:

Postulate 1: The state of a physical system is represented by the quantum state $|\Psi\rangle$ which belongs to a Hilbert space H.

Postulate 2: Every observable quantity $\mathcal{A}$ is represented by a Hermitian operator A acting in the Hilbert space H.

Postulate 3: The result of measuring on observable $\mathcal{A}$ must be one of the eigenvalues of the corresponding operator A.

Postulate 4: When an observable $\mathcal{A}$ is measured in the state $|\Psi\rangle$, the probability of obtaining the eigenvalue a_n of A is given by the squared amplitude of the corresponding eigenvector $|\Psi_n\rangle$. The state $|\Psi\rangle$ can be expanded in terms of $|\Phi_n\rangle$:

$$|\Psi\rangle = \sum_n c_n|\Phi_n\rangle \tag{2.1}$$

.

The probability of obtaining the eigenvalue a_n is therefore

$$P(a_n) = |c_n|^2 = |\langle\Phi_n|\Psi\rangle|^2 \tag{2.2}$$

.

Postulate 5: If the measurement of an observable $\mathcal{A}$ results in a_n then the state collapses to the normalized projection of $|\Psi\rangle$ onto the subspace associated with a_n

$$\Psi \rightarrow \frac{P_n|\Psi\rangle}{\sqrt{\langle\Psi|P_n|\Psi}} \tag{2.3}$$

Where P_n is the projection operator

$$P_n = \sum_m |\Phi_m\rangle\langle\Phi_m| \tag{2.4}$$

and m sums over all the eigenvectors which correspond with the eigenvector a_n.

Postulate 6: The state $|\Psi\rangle$ evolves in time according to the Schrödinger equation

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = H|\Psi(t)\rangle \tag{2.5}$$

2.1.2 The Schrödinger Equation

The Schrödinger equation is a differential equation which forms the basics for most quantum mechanical computations. It is often helpful to separate the Schrödinger equation into two pieces, the time-dependent component and the time-independent component. To do so, we must assume that the state $|\Psi(x,t)\rangle$ can be separated into a component depending entirely on position, $|X(t)\rangle$, and a component depending entirely on time, $|T(t)\rangle$, so that

$$|\Psi(x,t)\rangle = |X(x)\rangle|T(t)\rangle. \tag{2.6}$$

When we then apply this assumption to the Schrödinger equation, we see

$$i\hbar\frac{d}{dt}|X(x)\rangle|T(t)\rangle = H|X(x)\rangle|T(t)\rangle \tag{2.7}$$

We can now separate these expressions so each side is dependent on only one variable. We can then equate them to a constant which we will call E.

$$|T(t)\rangle^{-1}i\hbar\frac{d}{dt}|T(t)\rangle = |X(x)\rangle^{-1}H|X(x)\rangle = E \tag{2.8}$$

This separation allows us to write each side as an ordinary differential equation. Beginning with the time dependent component, we have

$$i\hbar \frac{d}{dt}|T(t)\rangle = E|T(t)\rangle. \tag{2.9}$$

This is a simple first order differential equation with a well known solution,

$$|T(t)\rangle = e^{-iEt}. \tag{2.10}$$

We have set $\hbar = 1$, a standard which will be convenient for future calculations. The details of the position-dependent calculation depend on the operator, H, but can be written generally as the time-independent Schrödinger equation:

$$H|X(x)\rangle = E|X(x)\rangle. \tag{2.11}$$

This takes the form of a simple eigenvalue problem, where the eigenvalues of the Hamiltonian, H, are the energy levels. Assuming we have a time-independent solution, $|X(x)\rangle$, the full time-dependent solution is

$$|\Psi(x,t)\rangle = e^{-iEt}|X(x)\rangle = \sum_{n=0}^{\infty} \frac{(-it)^n E^n}{n!}|X(x)\rangle = \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} H^n|X(x)\rangle = U(t)|X(x)\rangle \tag{2.12}$$

Here we've defined the time-evolution operator, $U(t)$, as

$$U(t) = e^{-iHt}. \tag{2.13}$$

This time-evolution operator allows us to take a solution to the time-independent Schrödinger equation and evolve it in time to a final time t.

2.2 Numerical Techniques

So far, we have sketched the basic theory of quantum mechanics. This theory has focused on analytical manipulation, and as a result it may not be obvious how to actually calculate observables from a system of interest using this formalism. In this section, we describe the basic numerical techniques used to calculate energies, positions, or other observables of interest. We will begin with a discussion of linear algebra representation of states and operators and basic implementation of the numerical description of a system. We include the simple example of a one dimensional harmonic oscillator. We will then proceed with a discussion of reduced basis methods, which are used heavily in Chapter 3. We will then finish with a discussion of Monte-Carlo methods, which are used commonly in many-body physics but we will also use to improve the sampling performance of trimmed sampling in Chapter 3.

2.2.1 Bases and Vectors

When we wish to calculate an observable from a complicated quantum system, be it position, momentum, energy, or some other parameter, we may turn to numerical simulations when a direct analytical solution eludes us. However, to perform numerical calculations of quantum systems, we must be able to represent those quantum systems in a way which is parseable by computers. In this subsection, we will discuss the basics or representing quantum systems in a numeric format and possible drawbacks of these representations. We will consider the quantum harmonic oscillator as a toy problem which we can treat numerically. To do so, we need to map the language of operators and functions onto that of matrices and vectors. The notation we use is Dirac notation, also called bra-ket notation.

2.2.1.1 Constructing Numerical Bases

We write a given state as the column vector, the 'bra' in 'bra-ket notation', $|\alpha\rangle$ where α represents all parameters required to define the state. In a continuous, infinite dimensional space, this is no different than defining the state as a function, $\psi(\alpha)$. The complex conjugate state, sometimes called the 'ket', is written as $\langle\alpha|$. We can define the inner product of two states, α and β, as $\langle\alpha|\beta\rangle$. In a continuous space, this takes the form of an integral,

$$\langle\alpha|\beta\rangle = \int \psi(\alpha)^\dagger \psi(\beta) d\Omega \tag{2.14}$$

where $\psi(\alpha)^\dagger$ signifies the complex conjugate of $\psi(\alpha)$ and $d\Omega$ signifies an integration over the full complex space wherein the vectors lie. However, for a finite dimensional space, this integral becomes a sum,

$$\langle\alpha|\beta\rangle = \sum_k \psi_k(\alpha)^\dagger \psi_k(\beta) \tag{2.15}$$

where now k is the index for the finite dimensional vectors and ψ_k is the kth element of ψ. We see immediately that this is just the well-known dot product of two vectors.

Since we may treat these states as complex vectors, we can think of them as a linear combination of basis vectors,

$$|\psi\rangle = \sum_n c_n |n\rangle \tag{2.16}$$

where n labels the basis vectors and $c_n = \langle n|\psi\rangle$, with the assumption that the basis is orthonormal, meaning $\langle n|m\rangle = \delta_{nm}$ where n and m label two basis vectors and δ is the Kronecker delta.

This machinery allows us to represent states as vectors. In many cases, such as the position of a particle, such a vector may require infinite dimensions. However, we

can often instead discretize the continuous variable, allowing us to represent it in finite dimensions. We will do exactly that in the quantum harmonic oscillator example.

First, however, we must be able to represent operators in this same language. The effect of operators in general is to map one vector in the Hilbert space onto another:

$$\hat{O}|\alpha\rangle = O_{\alpha\beta}|\beta\rangle. \tag{2.17}$$

where $O_{\alpha\beta}$ is some constant which depends on α and β. This is the same effect that matrices have when multiplying vectors in linear algebra. If we multiply $\langle\beta|$ on each side, and assume $\langle\beta|\beta\rangle = 1$, meaning the state is normalized, we get

$$\langle\beta|\hat{O}|\alpha\rangle = O_{\alpha\beta}. \tag{2.18}$$

We call $O_{\alpha\beta}$ the matrix element of $\hat{O}$. We will want to represent these matrix elements in terms of our basis vectors, O_{nm}. Doing so allows us to represent operators in our finite dimensional space as matrices, whose elements are defined as $O_{nm} = \langle n|O|m\rangle$.

2.2.1.2 Harmonic Oscillator

For a concrete example, let us consider the quantum harmonic oscillator in one dimension. We define the Hamiltonian for this system as

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{2.19}$$

where m is the mass of the particle and $\omega = \sqrt{k/m}$ is the frequency of the oscillation, with oscillator strength related to the spring constant k. We write $\hat{p}$ as the operator whose eigenvalues corresponds to the momentum of the particle and $\hat{x}$ as the operator whose eigenvalue correspond to the position of the particle. Note that, at this stage, these operators work in a continuous and therefore infinite-dimensional space.

Using the Schrödinger equation, we have a simple eigenvalue problem

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{2.20}$$

where we have used ψ to denote the state of a particle. While this is a well known problem and analytically solvable, we wish to represent it on a computer's architecture. First, we must discretize and limit the space we are considering. We will consider the position space representation of our particles. This means that our basis states $|n\rangle$, will be the box functions of width Δx and height $\frac{1}{\sqrt{\Delta x}}$ and positioned with their left edges at $x = n\Delta x$. We define $\Delta x = L/k$, where L is the length of the space we are considering and k is the number of basis states we will consider.

It is easy to see that this basis is orthonormal. The integral can only be non-zero if both box functions overlap, which only happens when $n = m$. When they do so, we integrate the box function from $n\Delta x$ to $(n+1)\Delta x$, with height $\frac{1}{\Delta x}$, so $\langle n|m\rangle = 1$.

We now consider the operators which define our Hamiltonian. The position operator, $\hat{k}$, has position as its eigenvectors. Meaning if we consider the matrix elements of $\hat{x}$ in our basis, we see

$$\langle m|\hat{x}|n\rangle = n\Delta x\langle m|n\rangle = n\Delta x\delta_{nm}. \tag{2.21}$$

In our position representation, the $\hat{x}$ operator is a diagonal matrix, with diagonal entries which correspond to the position of those basis vectors. To calculate the matrix elements of the momentum operator, we make use of our knowledge that in the position space representation, the momentum operator behaves as proportional to the differntial operator, $\hat{p} = -i\hbar\frac{d}{dx}$. We continue to use $\hbar = 1$, so the operator, when squared, is the second derivative:

$$\hat{p}^2 = -\frac{d^2}{dx^2}. \tag{2.22}$$

When discretized, this differential operator becomes the finite difference of the position wavefunction. In most cases, the first order finite difference is more than sufficient:

$$-\frac{d^2}{dx^2}\psi(x) \approx \frac{-\psi(x+\Delta x) + 2\psi(x) - \psi(x-\Delta x)}{\Delta x^2} \tag{2.23}$$

We apply this to our discrete case to find

$$\hat{p}^2|n\rangle = \frac{1}{\Delta x^2}\left(-|n+1\rangle + 2|n\rangle - |n-1\rangle\right) \tag{2.24}$$

and finally multiplying by $\langle m|$ to get the matrix elements, we find the equation for the matrix elements of the finite difference matrix,

$$\langle m|\hat{p}^2|n\rangle = (p^2)_{nm} = \frac{1}{\Delta x^2}\left(-\delta_{n+1,m} + 2\delta_{nm} - \delta_{n-1,m}\right) \tag{2.25}$$

In this representation, the finite difference matrix is the form of a Toeplitz matrix, with $\frac{2}{\Delta x^2}$ along the diagonal and $\frac{-1}{\Delta x^2}$ along each off-diagonal. Higher orders of finite difference may also be used when extra accuracy is needed.

We are now able to write an expression for the Hamiltonian matrix elements

$$H_{nm} = \frac{1}{2m}\frac{1}{\Delta x^2}\left(-\delta_{n+1,m} + 2\delta_{nm} - \delta_{n-1,m}\right) + \frac{1}{2}m\omega^2 n^2\Delta x^2\delta_{nm} \tag{2.26}$$

By following this definition, we can construct the matrix representation of the quantum harmonic oscillator Hamiltonian in the discretized position basis. The result is a $k \times k$ matrix. We show the result of diagonalizing such a matrix in figure 2.1. For this example, we took $\omega = m = 1$ and $k = 1000$. We diagonalize the resulting matrix to find its lowest eigenvalues and their corresponding eigenvectors, which we plot versus position. The calculated states match the well-known exact solutions to the 1d harmonic oscillator. With our choice of $\omega = m = 1$, we expect these states to correspond to eigenvalues 0.5, 1.5, and 2.5, and this is indeed what we find.
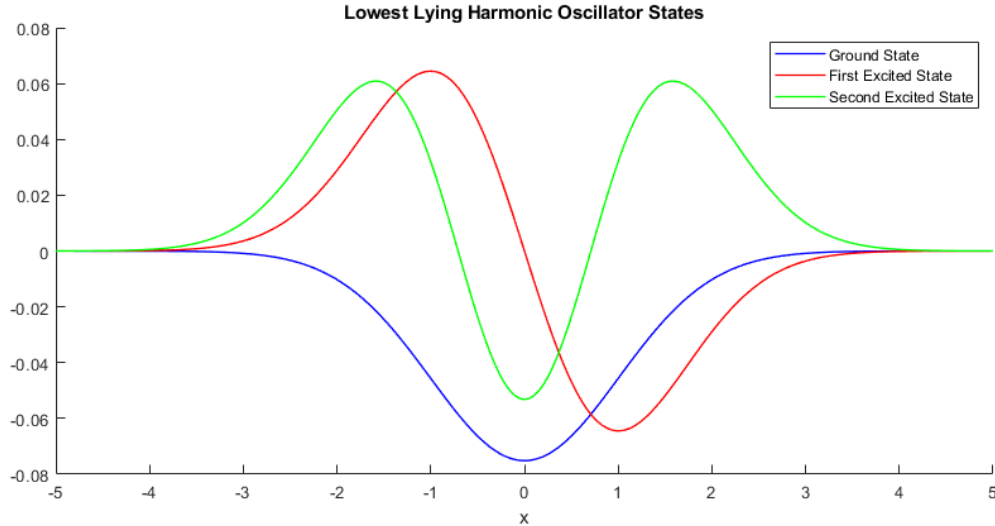
Figure 2.1: Lowest lying harmonic oscillator wavefunctions calculated in position basis. We used a $k = 1000$ lattice to discretize position, with $\omega = m = 1$. We diagonalized the resulting Hamiltonian matrix and look at the eigenvectors corresponding with the lowest three eigenvalues to see the well known harmonic oscillator wavefunctions corresponding to the ground state, first excited state, and second excited state.

For this simple example, we chose a fairly large value for k, and so our basis was constructed from 1000 position basis vectors. We did so because the effects of discretization diminish as Δx decreases. However, this comes at the cost of more memory and processor usage. The choice of basis plays a large roll in the effect of discretization errors, computational scaling, and overall accuracy.

2.2.1.3   Nonorthogonal Bases

In the previoius sections, we assumed that our basis is orthonormal, which is to say that $\langle n|m \rangle = \delta_{nm}$. However, as we will see in future sections, we may encounter situations where this assumption no longer holds. In such a case, we construct the norm matrix, also called the Gram matrix, for the space by considering the matrix element $N_{nm} = \langle n|m \rangle$. For an orthonormal space, the norm matrix is simply the identity matrix. However, in a nonorthonormal space the norm matrix is a symmetric, positive definite matrix.

For such a space, the above procedures still hold, however rather than solving a standard eigenvalue problem to calculate the energies and eigenfunctions of the Hamiltonian, or another operator, we instead solve the generalized eigenvalue problem:

$$H|\psi\rangle = NE|\psi\rangle \tag{2.27}$$

where now N is the norm matrix. States are normalized with respect to this norm matrix, so $\langle \psi|N|\psi \rangle = 1$. The process to solve this eigenvalue problem is usually fairly simple.

We multiply both sides of equation 2.27 by the inverse of the norm matrix, N^{-1}. This results in

$$N^{-1}H|\psi\rangle = E|\psi\rangle. \tag{2.28}$$

We multiply the Hamiltonian by the inverse of the norm matrix on the left. The resulting eigenvalue problem can be solved as a typical eigenvalue problem. However, during this process the norm matrix is inverted. As a result, if the norm matrix is poorly-conditioned, meaning it has eigenvalues very close to 0, then the resulting inverted matrix can be extremely sensitive to noise. This is often the case when basis vectors are very similar so that $\langle n|N|m\rangle \approx 1$ for $n \neq m$, resulting in significant off-diagonal elements. This is the central problem we address in Chapter 3.

2.2.2   Reduced Basis Methods

The space of most many-body systems grows exponentially as we increase the system size. While in some cases, such as the Lipkin model, symmetries allow us to reduce this to a more manageable basis, in many cases this is not possible, or the possible reduction in dimensionality is still insufficient to make the problem reasonable to calculate. To address this shortcoming, we can use various Reduced Basis methods to find a smaller and more tractable basis which still contains the physics we are interested in. The general prescription for this will be to find a set of states which are not degenerate with each other but which each have some overlap with the state we are interested in. We then project the Hamiltonian and any observables we are interested in onto this reduced basis, which allows us to carry out our computation in the smaller and more tractable basis. The particulars of determining the states used for this reduced basis are specific to each method, and the best method to use depends on the problem at hand.

In this section, we will look at three reduced basis methods: eigenvector continuation, Euclidean time projection, and generator coordinate method. In Eigenvector continuation, we project the Hamiltonian onto a subspace formed by the ground state of similar Hamiltonians. In Euclidean time projection, we form our reduced basis from various time evolutions of a sample starting state. In generator coordinate method, we will form a reduced basis from a parameterized ansatz.

2.2.2.1   Eigenvector Continuation

Eigenvector continuation is a method which finds the extremal eigenvalues and eigenvectors of a Hamiltonian which depends on one or more control parameters. It was first introduced in (2, 3) and further developed in (4, 5). Let us say we have a Hamiltonian which we can write as depending on a set of parameters, c_i. The core idea of eigenvector continuation is that as we vary those control parameters, the extremal eigenvectors also vary; however, the space in which those eigenvectors vary is substantially smaller than

the full space of the system. This implies that if we can project the full Hamiltonian onto the smaller space wherein the extremal eigenvectors lie, we are able to capture the interesting behavior of the system in a much smaller and thus less expensive subspace.

The typical prescription for eigenvector continuation is to build a reduced basis from the training vectors from some chosen training values of the control parameters. We then only need to solve the full system for a handful of values of the coupling, which can be done in an easier and more tractable sector, then project the system onto the subspace of those vectors. Because these eigenvectors are typically not orthonormal, this results in a generalized eigenvalue problem which we can then solve to find the extremal eigenvalues.

Let us consider a simple one-parameter family of Hamiltonian matrices defined as $H(c) = H_0 + cH_1$ where both H_0 and H_1 are finite-dimensional Hamiltonian matrices, and c is a parameter which we can control. We will define $|v(c)\rangle$ as the ground state of the Hamiltonian $H(c)$. For this given family of Hamiltonians, $H(c)$, we are interested in finding the ground state for a certain target coupling, which is to say we are trying to find $|v(c_t)\rangle$ where c_t is the target coupling. More formally, we are interested in solving the eigenvalue problem

$$H(c_t)|v(c_t)\rangle = E(c_t)|v(c_t)\rangle \tag{2.29}$$

for the lowest eigenvalue, $E(c_t)$, and its associated eigenvector. The dimension of this eigenvalue problem can be quite substantial, and may require numerical techniques to find the eigenvalues and eigenvectors if it is possible at all for the target coupling.

It may be the case that for our target value of the coupling, c_t, we are not able to diagonalize the Hamiltonian directly. It may also be the case that we are interested in many values of the coupling and the cost to diagonalize the Hamiltonian for each value of the coupling is too costly. For example, it may be the case that we can calculate the eigenvectors and eigenvalues of $H(c)$ only when c is small, but we are interested in larger values of c. We will take k values of the coupling, $c_1 = \{c_1, c_2, \cdots, c_k\}$, for which we are able to calculate the extremal eigenvectors of the Hamiltonian, $|v(c_i)\rangle$.

We now construct a new basis from this set of vectors, $|v(c_1)\rangle, \cdots, |v(c_k)\rangle$. We are interested in the ground state eigenvalue and eigenvector of the Hamiltonian at the target coupling, $H(c_t)$. We define the projected Hamiltonian operator $\tilde{H}$ as

$$\tilde{H}_{ij} = \langle v(c_i)|H(c_t)|v(c_j)\rangle. \tag{2.30}$$

The eigenvectors which form our basis, $|v(c_i)\rangle$, need not be orthonormal to each other. In fact, it is highly unlikely that they are, as this would indicate a discontinuity in the eigenvectors as the parameter c is varied. As a result, we also need the norm matrix, also called the Gram matrix:

$$N_{ij} = \langle v(c_i)|v(c_j)\rangle. \tag{2.31}$$

We can now write the problem in terms of our reduced basis as a generalized eigenvalue problem of the form

$$\tilde{H}|w(c_t)\rangle = \tilde{E}N|w(c_t)\rangle. \tag{2.32}$$

The dimension of this generalized eigenvalue problem is now equal to k, and is therefore easily diagonalizable. The ground state of this reduced system, $\tilde{E}$, is a good approximation for the ground state energy of the full system at the target coupling. Similarly, a good approximation for the target eigenvector, $|v(c_t)\rangle$, is $|v\rangle_{EC} = T_V|w(c_t)\rangle$ where T_V is the matrix of training vectors, returning the vector in the reduced basis back to the full Hilbert space:

$$T_V = [|v(c_1)\rangle; |v(c_2)\rangle \cdots |v(c_k)\rangle] \tag{2.33}$$

where the vectors are column vectors. The accuracy of these approximations depends on the training points taken. Training points which behave more similar to the ground state at the target coupling will likely provide a better approximation. Including more training points can likewise improve the accuracy, but can do so at the cost of causing the norm matrix to become ill-conditioned. This is a problem which we address using trimmed sampling in Chapter 3.

This process can be extended in a number of ways. If we are interested in excited states of the Hamiltonian, we must include excited states for each of the target couplings into the reduced basis. Otherwise, the process proceeds in the same manner. As the excited states also trace a trajectory through a lower dimensional subspace of the Hilbert space, their inclusion allows the reduced basis to also include these excited states.

Eigenvector continuation can also easily be used to calculate extremal eigenvalues and eigenvectors in families of Hamiltonians which depend on more than one parameter. To do so, we include sample points which vary along all parameters. This may require a much larger subspace if the number of parameters is large, but the particulars are not important for the remainder of this work.

As we will see in the example in Chapter 3, eigenvector continuation can be extremely sensitive to noise. As the training vectors, $|V(c_i)\rangle$, are often quite similar, the resulting norm matrix is often very poorly conditioned. As we discussed previously, when solving the generalized eigenvalue problem in an ill-conditioned basis, even tiny errors can overwhelm the calculation. As a result, eigenvector continuation can become very noise-sensitive when training points are taken close together or if there are many of them. We discuss how to address these problems in Chapter 3.

In summary, eigenvector continuation is a powerful method to calculate the extremal eigenvalue of a Hamiltonian which resides in a very large space. We need only be able to calculate the extremal eigenvectors of the system of interest for a few training values of the control parameter. This can be accomplished directly, through any number of iterative techniques, or through Monte-Carlo sampling. Regardless of the method to obtain these

vectors, we can then use them to construct a reduced basis onto which we project the Hamiltonian of the target value of the control parameter. We can then solve the resulting generalized eigenvalue problem to calculate a good approximation of the eigenvalue and eigenvectors of the system in the full Hilbert space.

2.2.2.2   Euclidean Time Projection

Euclidean time projection is another reduced basis method which can be used to project a complicated Hamiltonian into a much smaller subspace. For euclidean time projection, the basis we are projecting onto is formed by evolving an initial state $|v_0\rangle$ by using the Euclidean time projection operator, $e^{-H\tau}$. This is also sometimes called the imaginary time projection operator and can be seen as evolving a state in complex time. The exponential decay reduces the overlap of high energy states faster than the low energy states, and so the resulting basis states have higher overlap with the ground state and other low energy states once they have been projected further through imaginary time.

In particular, if we have some target ground state $|psi_{\mathrm{gs}}\rangle$, we can write the central idea as

$$|\psi_{\mathrm{gs}}\rangle = \frac{\lim_{\tau\to\infty} e^{-\tau H}|\psi_0\rangle}{\lim_{\tau\to\infty} e^{-E_0\tau}} \tag{2.34}$$

where we have some initial state $|\psi_0\rangle$ which has nonzero overlap with the ground state. While numerical methods exist for the exponentiation of matrices, in practice we use the Trotter approximation (6)

$$e^{-\tau H} = \prod_{n=1}^{N} e^{-\Delta\tau H} \approx \prod_{n=1}^{N} (1 - H\Delta\tau) \tag{2.35}$$

where $\Delta\tau = \frac{\tau}{N}$. We can thus iteratively take incremental steps along the time evolution, normalizing the state at each step

$$|\psi\rangle_{i+1} = \frac{(1 - H\Delta\tau)|\psi\rangle_i}{|(1 - H\Delta\tau)|\psi\rangle_i|} \tag{2.36}$$

until we reach our target time τ. Rather than simply iterating repeatedly, we can instead use some of these states as our basis. Each state should have progressively better overlap with the ground state as we evolve them in imaginary time.

We take a series of imaginary time slices, τ_n, and evolve the initial basis vector

$$|v_n\rangle = e^{-H\tau}|v_0\rangle \tag{2.37}$$

using the Trotter approximation if we wish until we reach the target time. We can take as many basis vectors as we wish, then write the projected Hamiltonian in this basis as

$$\tilde{H}_{nm} = \langle v_n|H|v_m\rangle \tag{2.38}$$

We know for certain that these states are not orthogonal, as we could not evolve them in imaginary time from other basis states if they were. As a result, we need to construct the norm matrix,

$$N_{nm} = \langle v_n | v_m \rangle \tag{2.39}$$

And we can now write the problem in terms of the reduced basis as a generalized eigenvalue problem

$$\tilde{H} | w \rangle = \tilde{E} N | w \rangle \tag{2.40}$$

where the dimension of the generalized eigenvalue problem is the number of time slices we took as our basis. The ground state energy of this generalized eigenvalue problem is a good approximation of the ground state energy of the full system, and we can calculate the approximation for the exact ground state vector as $| v_{\text{gs}} \rangle = T_V | w \rangle$, where T_V is the matrix of basis vectors used to return the reduced basis vector to the full vector space

$$T_V = [| v_0 \rangle; | v_1 \rangle \cdots | v_k \rangle] \tag{2.41}$$

where the vectors are column vectors. The accuracy of this method depends on the initial state and requires it to have nonzero overlap with the ground state. As more basis states are added, and as the length of the time projection increases, Euclidean time projection becomes more accurate, but increasingly sensitive to noise. At large imaginary time, each iteration of the time step produces less and less change to the resulting vector, and so the norm matrix can easily become very badly degenerate.

Euclidean time projection is a method to calculate the low lying energy levels of a system with a Hamiltonian which resides in a very large space. We take iterative time slices by evolving some initial state in imaginary time and use the resulting states to construct a reduced basis. The resulting generalized eigenvalue problem can then be solved for the ground state and other low lying states without the need to solve the eigenvalue problem in the full Hilbert space.

2.2.2.3 Generator Coordinate Method

In the generator coordinate method (GCM) we will construct a basis from a family of generating functions. These generating function scan depend on many parameters but span the Hilbert space we wish to understand. The central idea is that we can include a large number of these basis functions to produce a highly accurate reduced basis which does not require any prior diagonalization of the Hamiltonian or extensive matrix multiplications. Instead, we can rely more on analytic computations to determine the matrix elements of the reduced basis Hamiltonian. We then solve its associated eigenvalue problem to calculate energies.

The basic ansatz for generator coordinate method is an integral over generating functions

$$|\Psi\rangle = \int d\alpha f(\alpha)|\alpha\rangle \tag{2.42}$$

where $|\alpha\rangle$ is the generating function, α represents all parameters of the generating function, and $f(\alpha)$ is an associated weight function. The form of the generating function depends on the problem specified, but if each generating function maintains a particular symmetry, then the state $|\Psi\rangle$ retains this symmetry.

The expectation value of the trial wavefunction $|\Psi\rangle$ is given by

$$E = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int d\alpha d\alpha' f(\alpha)f(\alpha')\langle\alpha|\hat{H}|\alpha'\rangle}{\int d\alpha d\alpha' f(\alpha)f(\alpha')\langle\alpha|\alpha'\rangle}. \tag{2.43}$$

We define $H(\alpha, \alpha') = \langle\alpha|\hat{H}|\alpha'\rangle$ and $N(\alpha, \alpha') = \langle\alpha|\alpha'\rangle$. To find the extremal eigenvalues, we need we minimize (or maximize) the energy with respect to the weight functions,

$$\frac{\partial E}{\partial f(\alpha)} = 0 = \frac{\int d\alpha' H(\alpha, \alpha')f(\alpha) - E \int d\alpha' N(\alpha, \alpha')f(\alpha')}{\int d\alpha' d\alpha f(\alpha)f(\alpha')N(\alpha, \alpha')} \tag{2.44}$$

we can then rearrange the numerator to obtain the Hill-Wheeler equations

$$\int d\alpha H(\alpha, \alpha')f(\alpha) = E \int d\alpha' N(\alpha, \alpha')f(\alpha). \tag{2.45}$$

These equations have the form of a generalized eigenvalue problem, $H|\alpha\rangle = EN|\alpha\rangle$. The most direct approach would be to discretize the parameters, α, allowing us to represent the Hill-Wheeler equations in matrix form. We can then invert the norm matrix and solve the eigenvalue problem. However, as the generating functions are often very similar for similar values of the parameters, the norm matrix is very poorly conditioned. This can cause significant errors when solving the eigenvalue problem this way.

Instead, various other methods exist to solve the resulting problem. One popular approach is Löwdin symmetric orthogonalization. We need to construct a new orthogonal basis by diagonalizing the norm matrix. We start out by finding a unitary matrix U which diagonalizes N

$$N_{\text{diag}} = U^\dagger N U. \tag{2.46}$$

Because N is constructed from the overlap of basis vectors, we can replace its elements with their square roots then solve the equation for $N^{\frac{1}{2}}$

$$N^{\frac{1}{2}} = U N_{\text{diag}}^{\frac{1}{2}} U^\dagger. \tag{2.47}$$

We can also calculate $N^{\frac{-1}{2}}$ in the same way

$$N^{\frac{-1}{2}} = U N_{\text{diag}}^{\frac{-1}{2}} U^\dagger. \tag{2.48}$$

where we have again calculated $N_{\text{diag}}^{\frac{-1}{2}}$ by taking the square root of its eigenvalues, and this time inverting them.

Now we define

$$g(\alpha) = N^{\frac{1}{2}} f(\alpha) \tag{2.49}$$

By applying this relationship to the generalized eigenvalue problem, we find

$$N^{-\frac{1}{2}} H N^{-\frac{1}{2}} g(\alpha) = E g(\alpha) \tag{2.50}$$

and we now have a standard eigenvalue problem. The implementation of symmetric orthogonalization is problem dependent, but can involve costly integration. However, it can be precise in some cases to the limits of the numerical integration used. For a specific example, see the section in Chapter 3.

Generator coordinate method is a powerful technique which allows us to construct a basis within which we are able to solve a very complicated many-body problem without every orthogonalizing the full Hilbert space Hamiltonian and can allow highly accurate simulations. However, its poorly conditioned basis can cause it to be sensitive to noise, as with the other reduced basis methods we have introduced here. In the case of generator coordinate method, symmetric orthogonalization is a useful tool to reduce the effects of noise, but at the cost of requiring costly numerical calculations. In some cases, the degeneracy may still be too bad for symmetric orthogonalization to repair at double precision, in which case the method will fail. We introduce trimmed sampling to solve these problems in Chapter 3.

2.2.3   Monte-Carlo Methods

Monte Carlo algorithms are a class of numerical techniques which rely on repeatedly generating random numbers for use in statistical sampling. One of the primary uses of Monte Carlo sampling is to numerically integrate an arbitarary function. Say we have some function of one variable, $f(x)$. We can calculate its integral over a closed interval $[ab]$,

$$\int_a^b f(x) dx. \tag{2.51}$$

If we want to numerically integrate this function, we can use a simple Riemann sum

$$I = \int_a^b f(x) dx \approx \frac{b-a}{N} \sum_{j=1}^N f(x_j) \tag{2.52}$$

where x_j's are at regularly spaced intervals. This method is extremely simple and straight-forward, and can be fairly accurate when a large number of sample points are selected. However, generalizing to a large dimensional integral becomes very challenging. If we have N grid points in each dimension, then the total points for the whole grid is N^d,

where d is the number of dimensions. Clearly this becomes entirely impractical beyond a handful of dimensions, so we instead turn Monte Carlo methods.

We can think of this integral as the area under the curve of the function. We will use random sampling to estimate the value of the integral. We can consider a box which spans the domain and range of $f(x)$. If we choose points randomly within this box, we can easily test if they are above or below the curve of $f(x)$. After a large number of samples, we can then estimate the integral as

$$\int_a^b f(x)dx = \lim_{N \to \infty} \left(\frac{N_{\text{below}}}{N} \right) A_\Omega \tag{2.53}$$

where N is the total number of points sampled, N_{below} is the number of points that were below the curve, and A_Ω is the area of the box within which we are sampling. We are therefore able to calculate the integral of this function without requiring much knowledge at all about the function itself beyond the ability to evaluate it.

This naturally can be extended to arbitrarily many dimensions, where we now sample over the hypervolume that bounds the domain and range of the function. The prescription follows very simply, and allows the integration of functions which might otherwise be impossible to integrate.

We can of course make use of Monte Carlo methods to sample arbitrary probability distributions. If we have a continuous probability distribution $\mathcal{P}$ and an operator we wish to sample $\hat{O}$, we can calculate its expectation value over that distribution as

$$\langle \hat{O} \rangle = \int \mathcal{P}(x) \langle x | \hat{O} | x \rangle dx \tag{2.54}$$

where $\mathcal{P}(x)$ is the probability associated with state $|x\rangle$. By Monte Carlo sampling, we can calculate the value of this integral exactly as above.

2.2.4 Importance Sampling

In many applications, such as when calculating thermal averages in statistical mechanics or path integrals in Euclidean-time field theory, integrals must be computed over many degrees of freedom weighted by an exponential Boltzmann factor,

$$\langle M \rangle = \frac{\sum_C M(C) e^{-\beta E(C)}}{\sum_C e^{-\beta E(C)}} \tag{2.55}$$

Because configurations are weighted exponentially, only a very small handful of configurations contribute meaningfully to the distribution. Randomly sampling the full hypervolume is thus very inefficient, as most samples will not affect the overall integral significantly.

Instead, we will use importance sampling to prioritize selecting samples which are more important to the final result. Instead of taking a uniform sample across the full

hypervolume, we instead select configurations according to a distribution which assigns greater weight to important sectors. In the case of the Boltzmann factor, we would sample according to

$$p_\beta(C) = \frac{e^{-\beta E(C)}}{\sum_{C'} e^{-\beta E(C')}} \tag{2.56}$$

We then need only average the quantity for which we wish to calculate the average over the states taken by sampling this distribution. Even with relatively few samples, importance sampling is able to effectively estimate these high dimensional integrals. One drawback of importance sampling is the difficulty of drawing samples from an arbitrary distribution, particularly if that distribution is not easily inverted. To sample from a complicated distribution, we will make use of Markov chains in a technique called Markov-chain Monte Carlo.

2.2.4.1 Markov-Chain Monte Carlo

Markov-chain Monte Carlo is a Monte Carlo method which allows us to sample arbitrary distributions. Consider a chain of configurations labeled by the order they were selected in. We will call this label τ, and the probability that configuration A is selected at computation step is $P(A, \tau)$. If we have a selected configuration A at step τ, we define the probability that we select configuration B at step $\tau + 1$ as $W(A \to B)$.

As long as the Markov process is ergodic, then once the chain grows long enough we will reach an equilibrium distribution, $lim_{\tau \to \infty} P(C, \tau) \to p(C)$. Our goal then is for $p(C)$ to approximate a target distribution, $p_T(C)$. One way to ensure this is to apply a condition called detailed balance. Detailed balance requires that

$$W(A \to B)p_T(A) = W(B \to A)p_T(B) \tag{2.57}$$

for every pair of configurations A and B. After many computation steps, we reach an equilibrium distribution satisfying

$$\sum_{B \neq A} W(A \to B)p(A) = \sum_{B \neq A} W(B \to A)p(B) \tag{2.58}$$

When we compare this to equation 2.57, we conclude that $p(A) = P_T(A)$ for all configurations A, and we have successfully sampled the distribution.

In this work, we will achieve this detailed balance condition by utilizing the Metropolis algorithm. The approach is to always accept steps in the Markov chain which improve on some cost function, often the energy of a configuration, and to accept steps which do not improve this cost function with a probability determined by the difference between the value of the cost function for the previous sample and the subsequent sample. Specifically, we say

$$W(A \to B) = \begin{cases} e^{-\beta[E(B) - E(A)]} & E(B) > E(A) \\ 1 & E(B) \leq E(A) \end{cases} \tag{2.59}$$

where we have used $E(A)$ to mean the energy of configuration A, though in general it can be used as a generic cost function. The coefficient β follows from the Boltzmann statistics prescription which we have been following, however it can more generally be any scaling coefficient.

A method which can improve sampling and prevent the chain from becoming stuck in local minima is called simulated annealing. Taking again the analog from statistical physics, we will begin with a system where the temperature is very high, and so β is very low. We then cool the system gradually, allowing β to increase and the Metropolis sampling to become more selective for later samples. This allows the chain to be more likely to accept configurations early and less likely to accept later configurations that do not strictly improve the cost function.

Monte Carlo methods form an important foundation for many computational techniques. They allow us to estimate averages of large complicated systems over a wide range of possible configurations, as well as numerically integrate high dimensional problems without suffering from high dimensionality as much as most other numerical integration techniques.

2.3 Quantum Many-body Physics

The quantum many-body problem describes the behavior of systems where we have many quantum particles, each of which may be governed by single particle Hamiltonians, but which may also have a term in their Hamiltonians which causes these particles to interact with each other. The resulting possibility space can grow exponentially with the number of particles, stretching modern computational techniques to their limits. In this section, we will describe the basic formalism needed to describe these systems.

Up to this point, we have described the quantum mechanics formalism for a single particle. We wish to describe a system involving many such quantum particles. We first define the one-body Hilbert space as $\mathcal{H}_1$. This is the space within which we can fully define solutions to the single particle Hamiltonian. For example, if a particle lives in a three-dimensional Euclidean space, we take $\mathcal{H}_1 = L^2(\mathbb{R}^3)$ with the Lebesgue measure.

Next, if we wish to describe the many-particle state, we use the composition of two or more such single-particle Hilbert spaces, $\mathcal{H}_A$ and $\mathcal{H}_B$. The composition of two Hilbert spaces is described by their tensor product, $\mathcal{H}_A \otimes \mathcal{H}_B$. If we have N particles, we describe the composite space of these particles as $\mathcal{H}_1^{\otimes N}$, where we use the notation that

$$\mathcal{H}_A^{\otimes N} = \underbrace{\mathcal{H}_A \otimes \cdots \otimes \mathcal{H}_A}_{N \text{ times}} \tag{2.60}$$

To calculate the many-body Hilbert space, we now take into account that these particles are indistinguishable. We must project $\mathcal{H}_1^{otimesN}$ onto the subspace which is appro-

priate for the particle statistics. In nearly all calculations, we are interested in Fermion or Boson statistics. For Fermions, we project onto the antisymmetric Hilbert space,

$$\mathcal{H}_N^{\text{fer}} = \mathcal{A}(\mathcal{H}_1^{\otimes N}) \tag{2.61}$$

and for Bosons we project onto symmetric Hilbert space

$$\mathcal{H}_N^{\text{bos}} = \mathcal{S}(\mathcal{H}_1^{\otimes N}) \tag{2.62}$$

Finally, we can represent the many-body Hamiltonian in this composite Hilbert space. As this Hamiltonian lives in the composite space of all particles, representations of it can become enormously expensive even for relatively low particle numbers. For even a very simple system, such as a lattice of N spin $1/2$ particles, the size of the Hilbert space has 2^N complex dimensions. Because the lattice grows exponentially with N, and many realistic physical systems include hundreds, thousands, or more particles, representing the full space on classical computer architecture quickly becomes impractical. Additional problems emerge from correlations caused by the interactions between particles. These correlations can be highly non-local and their effects are often difficult to quantify.

From the earliest days of quantum mechanics, many methods have been developed in an attempt to solve or circumvent these problems. They range from direct and straight-forward approaches to much more advanced and complex methods. We present here several important approaches which are commonly used. This is not meant to be an exhaustive list of methods, nor is it meant to be a full description of each method listed, but is rather a brief overview of the method.

2.3.1   Full Configuration Representation

In some cases, the problem is simple enough that it can be directly diagonalized. In such a case, we attempt to represent the full many-body Hilbert space Hamiltonian as a matrix and diagonalize it directly. This option is really only available in the simplest of models and even then typically only at very small scales. An example of this is the Lipkin model, which we will present later, with low particle number and system size, where the number of possible states can be relatively low. However, even in such simple models, it is often preferable to use faster and less memory-intensive methods, which we present in further sections.

2.3.2   Symmetry Method

By taking advantage of symmetries, we can stretch the full representation much further than might otherwise be possible. While the full-configuration representation is applicable to any problem we might be interested in, it fails to take advantage of the structure of

the specific problem. Often, the true space of possible configurations is much lower than it might appear when we consider symmetries of the problem. This will allow us to diagonalize several smaller matrices, rather than one immense one.

To begin, we identify any symmetries of the Hamiltonian and call them S_k. Because these symmetries commute with the Hamiltonian, $[S_k, H] = 0$, they share the set of eigenstates with the Hamiltonian. Say that S_k is an element of a symmetric group S which commutes with the Hamiltonian. We will call its eigenvalues s_n, such that $S_k|\psi_n\rangle = s_n|\psi_n\rangle$. We can write the matrix element of the Hamiltonian with respect to two states ψ_m and ψ_n as

$$H_{nm} = \langle\psi_n|H|\psi_m\rangle \tag{2.63}$$

We can now substitute $|\psi_m\rangle = \frac{1}{s_m}S_k|\psi_m\rangle$ to obtain

$$H_{nm} = \frac{1}{s_m}\langle\psi_n|HS_k|\psi_m\rangle = \frac{1}{s_m}\langle\psi_n|S_kH|\psi_m\rangle = \frac{s_n}{s_m}\langle\psi_n|H|\psi_m\rangle \tag{2.64}$$

Rearranging slightly,

$$s_mH_{nm} = s_nH_{nm} \Rightarrow (s_m - s_n)H_{nm} = 0 \tag{2.65}$$

If both $|\psi_n\rangle$ and $|\psi_m\rangle$ do not share the same symmetry, that is $s_n \neq s_m$, the H_{nm} must equal 0. As a result, we can write the Hamiltonian matrix as a block diagonal matrix where each block represents an eigenvalue of this symmetry.

The symmetry method can be a powerful tool to solve many-body problems, but as with the full configuration representation it can still suffer from rapid scaling as the number of particles increases. While many methods exist which allow us to calculate larger and more complicated systems, understanding the symmetry of the problem is always a useful tool to save time and effort.

2.3.3 Hartree-Fock Theory

The Hartree-Fock method approximates the ground state of a system by treating the system as a collection of non-interacting particles with a mean-field potential which emulates the interaction between particles. It neglects correlation effects between particles and the particles move individually in an effective one-body potential. We will vary the single-particle basis of a slater determinant $|\Phi\rangle$ to minimize the Hartree-Fock energy E_{HF} defined as

$$E_{\mathrm{HF}} = \min_{|\Phi\rangle} \frac{\langle\Phi|H|\Phi\rangle}{\langle\Phi|\Phi\rangle}. \tag{2.66}$$

Because Hartree-Fock is a variational method, the Hartree-Fock energy $E_{\mathrm{HF}} \leq E_0$, where E_0 is the ground state energy of the full Hamiltonian. This minimization problem produces a set of equations which can be solved to produce the Hartree-Fock Hamiltonian, whose eigenvalue problem may then be solved.

To derive the Hartree-Fock equations, we write a trial wavefunction in second quantization,

$$|\text{HF}\rangle = |\phi_1, \cdots, \phi_A\rangle = \prod_{i=1}^{A} \hat{a}_i^\dagger |0\rangle \tag{2.67}$$

where $\{|\phi_i\rangle\}$ is the set of single particle Hartree-Fock states with corresponding operators $\{\hat{a}_i^\dagger\}$. We can expand the single particle states and creation operators with respect to a fixed basis, often the harmonic oscillator basis

$$|\phi_i\rangle = \sum_p D_{ip}|\chi_p\rangle \tag{2.68}$$

$$\hat{a}_i^\dagger = \sum_p D_{ip}\hat{c}_a^\dagger \tag{2.69}$$

where $D_{ia} = \langle \chi_a|\phi_i\rangle$ is the matrix element of the unitary transformation, $|\chi_a\rangle$ is the single particle states, and $\hat{c}_a^\dagger$ is the creation operator for the reference basis.

We define the one-body density matrix of the trial state with respect to the reference basis as

$$\rho_{pq}^{(1)} = \langle \text{HF}|\hat{c}_p^\dagger \hat{c}_p|\text{HF}\rangle = \sum_{ij} D_{pi}D_{qj}^*\langle \text{HF}|\hat{a}_j^\dagger \hat{a}_i|\text{HF}\rangle = \sum_i^A D_{ip}D_{ji}^* \tag{2.70}$$

We can further define the corresponding density operator as

$$\hat{\rho} = \sum_{pq} \rho_{pq}^{(1)}|\chi_p\rangle\langle\chi_q| \tag{2.71}$$

After variation is performed, it is important to ensure that the many-body state is still a single Slater determinant. This is accomplished by requiring that the one-body density matrix is idempotent and hermitian

$$\sum_r \rho_{pr}^{(1)}\rho_{rq}^{(1)} = \rho_{pq}^{(1)} \tag{2.72}$$

$$\rho)pq^{(1)*} = \rho_{qp}^{(1)}$$

Let us say that we have a general two-body Hamiltonian in second quantized form

$$\hat{H} = \sum_{pq} t_{pq}\hat{c}_p^\dagger \hat{c}_q + \frac{1}{4}\sum_{pqrs} v_{pqrs}\hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r \tag{2.73}$$

where t_{pq} is the matrix elements of the kinetic energy operator and v_{pqrs} is the antisymmetric matrix elemetns of the two-body potential. The state $|\text{HF}\rangle$ is normalized, so the energy is

$$E = \langle \text{HF}|\hat{H}|\text{HF}\rangle = \tag{2.74}$$

$$\sum_{pq} t_{pq}\langle \text{HF}|\hat{c}_p^\dagger \hat{c}_q|\text{HF}\rangle + \frac{1}{4}\sum_{pqrs} v_{pqrs}\langle \text{HF}|\hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r|\text{HF}\rangle$$

We can also introduce the two-body density matrix

$$\rho^{(2)}_{pqrs} = \langle \text{HF} | \hat{c}^\dagger_p \hat{c}^\dagger_q \hat{c}_s \hat{c}_r | \text{HF} \rangle \tag{2.75}$$

which allows us to rewrite equation 2.74 in a more simple form

$$E = \sum_{pq} t_{pq} \rho^{(1)}_{pq} + \frac{1}{4} \sum_{pqrs} v_{pqrs} \rho^{(2)}_{pqrs}. \tag{2.76}$$

For a single Slater determinant, we can factor the two-body density into its one body components,

$$\rho^{(2)}_{pqrs} = \rho^{(1)}_{pq} \rho^{(1)}_{qr} - \rho^{(1)}_{pr} \rho^{(1)}_{qs}. \tag{2.77}$$

We can rewrite equation 2.76 in terms of one body density matrices,

$$E = \sum_{pq} t_{pq} \rho^{(1)}_{pq} + \frac{1}{2} \sum_{pqrs} v_{pqrs} \rho^{(1)}_{ps} \rho^{(1)}_{qr} \tag{2.78}$$

We can now perform a variation of the above term, neglecting terms which are quadratic or higher in $\delta \rho^{(1)}$.

$$\delta E = \sum_{pq} t_{pq} \delta \rho^{(1)}_{qp} + \frac{1}{2} \sum_{pqrs} v_{pqrs} (\delta \rho^{(1)}_{ps} \rho^{(1)}_{qr} - \rho^{(1)}_{ps} \delta \rho^{(1)}_{qr}) \tag{2.79}$$

We define an auxiliary mean-field Hamiltonian

$$h_{pq} = t_{pq} + u_{pq} \tag{2.80}$$

$$u_{pq} = \sum_{rs} v_{pqrs} \rho^{(1)}_{rs}$$

which, after some simplifying, allows us to write the stationary condition for the energy as

$$\delta E = \sum_{ik} h_{ik} \delta \rho^{(1)}_{ki} = 0 \tag{2.81}$$

From this, we know that the mean field Hamiltonian must commute wit the one body density operator,

$$[\hat{h}, \hat{\rho^{(1)}}] = 0, \tag{2.82}$$

so the mean-field Hamiltonian shares a common eigenbasis with the one body density operator. The eigenvalues of the mean field Hamiltonian, obtained from the eigenvalue problem

$$\hat{h} | \phi_n \rangle = \epsilon_n | \phi_n \rangle \tag{2.83}$$

are called the Hartree-Fock single particle energies. If we transform this eigenvalue problem to the reference basis,

$$\sum_r h_{pr} D_{ir} = \epsilon_i D_{ip} \tag{2.84}$$

the substitute the mean field Hamiltonian, we get

$$\sum_r (t_{pr} + \sum_q s \sum_j v_{pqrs} D_{js}^* D_{jq}) D_{ir} = \epsilon_i D_{ip} \tag{2.85}$$

This is called the Hartree-Fock equation and is a non-linear equation for the solutions to the eigensystem of $\hat{h}$. Solving this equation is typically done using iterative techniques.

2.3.4 Coupled Cluster Theory

The ansatz for coupled cluster theory is

$$|\Psi\rangle = e^T |\Phi_0\rangle \tag{2.86}$$

where $|\Phi_0\rangle$ is the reference state and we have defined the cluster operator T as

$$T = \sum_{p=n} T_n \tag{2.87}$$

where A is the maximum number of particle-hole excitations and

$$T_n = \left(\frac{1}{n!}\right) \sum_{i_1 \cdots i_n, a_1 \cdots a_n} t_{i_1 \cdots i_n}^{a_1 \cdots a_n} a_{a_1}^\dagger \cdots a_{a_n}^\dagger a_{i_n} \cdots a_{i_1} \tag{2.88}$$

We often truncate this operator to a small number of terms, for example we may truncate at $n = 2$, in what is called the singles and doubles approximation. We begin with the time-independent Schrödinger equation,

$$H|\Psi\rangle = E|\Psi\rangle \tag{2.89}$$

and insert the coupled cluster ansatz. Left multiplying by e^-T, then left multiplying by the reference state gives

$$E = \langle\Phi_0|\bar{H}|\Phi_0\rangle \tag{2.90}$$

while we can multiply by an excited state instead of the reference state to obtain

$$0 = \langle\Phi_{i_1 \cdots i_n}^{a_1 \cdots a_n}|\bar{H}|\Phi_0\rangle \tag{2.91}$$

. We have defined $\bar{H}$ as the similarity transformed Hamiltonian,

$$\bar{H} = e^{-T} H e^T \tag{2.92}$$

and the excited states are defined as

$$|\Phi_{i_1 \cdots i_n}^{a_1 \cdots a_n}\rangle = a_{a_1}^\dagger \cdots a_{a_n}^\dagger a_{i_1} \cdots a_{i_n} |\Phi_0\rangle \tag{2.93}$$

We define the reference energy as

$$E_{\text{ref}} = \langle\Phi|H|\Phi\rangle \tag{2.94}$$

and the correlation energy is the difference between the coupled cluster energy E and the reference energy,

$$\Delta E = E - \langle \Phi | H | \Phi \rangle \tag{2.95}$$

which subtracts off the reference energy. Coupled cluster theory can perform well for single determinant reference states with weakly correlated systems, but struggles with more strongly correlated systems.

3

Trimmed Sampling

3.1 Generalized Eigenvalue Problem

One method commonly used to find the extremal eigenvalues and eigenvectors of large quantum systems is to project the system into a reduced basis which contains the eigenvectors of interest. Because the methods used to project onto these bases do not in general produce an orthogonal, or even normalized, basis, this process results in a generalized eigenvalue problem with the form $H|\psi\rangle = EN|\psi\rangle$, where H is the projected Hamiltonian matrix, N is the norm matrix, $|\psi\rangle$ is the column vector of the projected eigenvector and E is the energy of that state. We can then further compute the expectation value of a given projected observable of interest, O, by using the formula $\langle O \rangle = \langle\psi|O|\psi\rangle/\langle\psi|N|\psi\rangle$.

Many methods exist which use generalized eigenvalue problems to reduce large systems to smaller, non-orthogonal bases. One well know technique in nuclear physics is the generator coordinate method, where the corresponding generalized eigenvalue problem is called the Hill-Wheeler equation(7, 8, 9, 10). It is used in several computational approaches utilizing variational subspace methods with non-orthogonal bases (11, 12, 13). Eigenvector continuation (2, 5, 14, 15, 16, 17) and the more general class of reduced basis methods(18, 19, 20) rely on this basis projection and solving the resulting generalized eigenvalue problem. It is also useful for Monte Carlo simulations where matrix elements are calculated from trial wavefunctions produced using Euclidean time projection from several different initial and final states(21, 22, 23, 24, 25).

The generalized eigenvalue problem allows a tractable solution to computing large quantum systems and can make efficient use of computational resources, but comes at the cost of a significant (and difficulty to quantify) sensitivity to noise. Solving the generalized eigenvalue problem relies on finding the eigenvectors and eigenvalues of $N^{-1}H$, and because the condition number of the norm matrix increases with the size of the subspace, even minor errors, such as limits of machine precision, can cause significant errors. The Hamiltonian may be likewise sensitive in some cases. This problem becomes even worse when using stochastic methods, such as Monte Carlo, where statistical errors

when calculating the elements of the Hamiltonian and norm matrices can cause very large errors when calculating observables. These challenges likewise arise for applications in quantum computing , where each measurement is stochastic in nature. Variational subspace methods in quantum computing (26) can produce both statistical and systematic errors due to gate errors , measurement errors, and decoherence. These resulting errors can cause any eigenvalue or observable calculated from the generalized eigenvalue problem to have substantial error(27, 28, 29, 30).

Some methods exist to address the errors when dealing with ill-posed problems. One popular approach is Tikhonov regularization (31) and its simplest implementation, ridge regression or nugget regularization. With ridge regression, a small multiple of the identity, ϵI, is added to the norm matrix before it is inverted. This has the effect of significantly stabilizing the norm matrix, but it is not straightforward to choose an appropriate value for ϵ (32), nor is it easy to estimate the systematic bias introduced by this regularization.

Trimmed sampling, first introduced in (33), is an approach which uses physics-based constraints and Bayesian interference (34, 35) to significantly reduce the errors of values calculated from the generalized eigenvalue problem. Rather than simply regulating the norm matrix, we sample the probability distributions of the norm and Hamiltonian matrix elements, as estimated from the method used to produce the generalized eigenvalue problem, then weight these distributions by a set of physics-informed likelihood functions, which may include any information knows about the problem. One such constraint is the positivity of the norm matrix, as by construction any reduced subspace must be positive definite. We can further apply constraints on the convergence of the energy eigenvalues or observables in question, or we an apply more problem-specific constraints as the opportunity arises.

In this chapter, we will describe the algorithm in greater detail, present several improvements which can be used to increase performance, and detail several examples showing this method and these improvements.

3.2   Bayesian Analysis

If we had exact calculations of the norm and Hamiltonian matrices, we would be able to calculate any observable within the reduced subspace without worrying about its sensitivity to noise. Unfortunately, we must assume that any calculation of the norm and Hamiltonian contains errors, even if these errors are on the order of machine precision. We can therefore consider our initial problem not as Hamiltonian and norm matrices with fixed elements, but rather as a distribution of Hamiltonian and norm matrices, with probability density determined by a prior estimation of uncertainty on the elements. We call this prior distribution function $P(H, N)$, and can calculate it by taking a product of

uncorrelated distributions of each element

$$P(H, N) = \prod_{ij} P(H_{ij})P(N_{ij}) \tag{3.1}$$

In most cases, we can assume Hamiltonian and norm elements' prior distribution is Gaussian with standard deviation ΔH_{ij} and ΔN_{ij} respectively, giving a prior distribution

$$P(H, N) = \prod_{ij} \frac{e^{-\frac{(H_{ij}-\tilde{H}_{ij})^2}{2(\Delta\tilde{H}_{ij})^2}} e^{-\frac{(N_{ij}-\tilde{N}_{ij})^2}{2(\Delta\tilde{N}_{ij})^2}}}{2\pi\Delta\tilde{H}_{ij}\Delta\tilde{N}_{ij}}. \tag{3.2}$$

With a more detailed model of the errors affecting the initial matrix elements, additional considerations such as correlations or asymetrical errors can be easily implemented.

When we calculate a value of interest, such as the ground state energy, any deviation from the exact calculation can produce substantial errors. For example, in some cases, these deviations can cause the norm matrix to lose positivity, and so the resulting system is physically unrealizable. As a result, we could immediately reject any results from this calculation–they are likely to be inaccurate. To generalize, we postulate a posterior likelihood function, $P(R|H, N)$, where R represents physics-informed criteria. In the example of norm matrix positivity, this likelihood function is simply a step function

$$P(R|H, N) = \begin{cases} 1, & N \text{ is positive definite} \\ 0, & \text{otherwise} \end{cases}$$

This likelihood function refers to rejecting any matrix that is not positive definite, and accepting any matrix which is. However, the likelihood function can include any information we may have about the problem. The examples in this chapter use a product of several such functions, which will be explained in those sections.

Using Bayes' theorem, we can calculate the posterior distribution, given by

$$P(H, N|R) = \frac{P(R|H, N)P(H, N)}{\int \prod_{ij}[dH_{ij}dN_{ij}]P(R|H, N)P(H, N)}. \tag{3.3}$$

This posterior distribution contains only physically realizable systems, weighted both by their similarity to the initially measured matrices and by our likelihood functions which help to further reject unstable and problematic systems. As a result, by sampling the value of interest, O, from the systems in this distribution, we get a final distribution of the observable and can calculate the average for a more accurate estimate. As an additional bonus, because we have this posterior distribution, we can also estimate the uncertainty of that observable.

3.3 Methods

In this section, we will describe a very basic implementation of Trimmed Sampling. While significant improvements to the algorithm can be achieved, this represents a simple demonstration of the method without the clutter of more advanced techniques.

Let us consider a pair of matrices, $\tilde{H}$ and $\tilde{N}$. For each matrix element in $\tilde{H}$ and $\tilde{N}$ have corresponding error estimates with standard deviations $\Delta\tilde{H}$ and $\Delta\tilde{N}$. These matrices form the generalized eigenvalue problem of interest, $\tilde{H}|\psi\rangle = E\tilde{N}|\psi\rangle$. We likewise have a matrix observable, $\tilde{O}$, from which we calculate a given observable, $\langle\tilde{O}\rangle$. In a straightforward calculation, we expect the value of $\langle\tilde{O}\rangle$ to vary from the exact value, especially as $\Delta\tilde{H}$ and $\Delta\tilde{H}$ become non-negligible.

Rather than simply calculate this observable, we instead add an additional term to $\tilde{N}$ and $\tilde{H}$ drawn from the distributions $P(\delta\tilde{H}_{ij}) = \frac{1}{\Delta\tilde{H}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\delta\tilde{H}_{ij}}{\Delta\tilde{H}}\right)^2}$ and $P(\delta\tilde{N}_{ij}) = \frac{1}{\Delta\tilde{N}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\delta\tilde{N}_{ij}}{\Delta\tilde{N}}\right)^2}$ where $\delta\tilde{H}_{ij}$ and $\delta\tilde{N}_{ij}$ are the matrix elements of the shifts added to $\tilde{H}$ and $\tilde{N}$ respectively. We call these the trial matrices, $\tilde{H}'$ and $\tilde{N}'$ and can then evaluate the likelihood function of these trial matrices, $P(R|H,N)$. In practice the evaluation of this likelihood function is highly dependent on the forms of the functions, so for simplicity we will present here a likelihood function which is applicable in many cases. This likelihood function takes the form

$$P(R|H,N) = \alpha f_{pos}(N) f_C(H,N). \tag{3.4}$$

The first term, α, is a normalization constant which cancels in 3.3. The second factor, $f_{pos}(N)$, is the requirement that the norm matrix be positive definite. It equals 1 in N is positive definite and 0 otherwise. The final factor, $f_C(H,N)$ is a function which relies on the submatrix energies E_n given by solving the generalized eigenvalue problem with only the first n basis states,

$$H^{(n\times n)}|\psi_n\rangle = E_n N^{(n\times n)}|\psi_n\rangle. \tag{3.5}$$

We assume that the basis states $|v_1\rangle, |v_2\rangle, \cdots, |v_n\rangle, \cdots$ have been ordered so that the energies E_n converge to $E_e xact$ as a smoothly-varying function of n for sufficiently large n. If the variational principle applies to the method we are using, such as with eigenvector continuation, then the sequence E_n must be monotonically decreasing and bounded from below. We define the convergence ratio C_n for $n > 2$ as

$$C_n = \frac{E_n - E_{n-1}}{E_{n-1} - E_{n-2}} \tag{3.6}$$

. This convergence ratio is the ratio of energy differences for consecutive energies E_n. Let C_{max} be the maximum of C_n over all n and C an estimate of the upper bound for C_{max} of an exact system. The second likelihood factor, $f_C(H,N)$ has the form

$$f_C(H,N) = e^{-\frac{C_{max}}{C}}. \tag{3.7}$$

This likelihood function penalizes the likelihood (and therefore the final weighting if an average is taken) for the trial matrices if the convergence rate for their ground state energies are much slower than expected or violates the variational principle. Neither the value of C nor the functional form of $f_C(H, N)$ are essential features and do not need to be fine tuned. Similar results can be obtained from a wide range of choices and will be discussed in a later section.

We calculate the observable, $\langle O \rangle_{trial}$ and store it along with the value of the likelihood function for the trial matrices $\tilde{H}'$ and $\tilde{N}'$, and repeat this sampling process a large number of times, until we have sufficiently sampled the prior distribution. Finally, we calculate the weighted median of each calculated observable, weighted by the likelihood function calculated with it. This importance reweighting scheme is commonly used in Markov Chain Monte Carlo algorithms (36), and is similar to some other sampling/importance reweighting methods used (37, 38), but differs in that we are not resampling the data.

The weighted median of the observable $\langle O \rangle$ represents an estimate of the exact value of the observable, and calculating the 14th and 86th percentiles can give a reasonable estimate of the width of the distribution. Taking the weighted median, rather than the weighted mean, ensures that the final estimate is a physical example, a condition not guaranteed if the Hilbert space of the physical system is not convex.

3.4   Two Parameter Example

For a simple, easily visualized example, consider a two state generalized eigenvalue problem of the form

$$N = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}, H = (x+y) \begin{bmatrix} 1 & xy \\ xy & 1 \end{bmatrix} \tag{3.8}$$

Where $r = \sqrt{x^2 + y^2}$. We can see that when $r < 1$, the Norm matrix is positive definite. As this model is arbitrary and for illustrative purposes, we will assume that we are interested in a bound state of this system and that the gap between the ground state and the excited state is approximately 3. We can plot this two parameter problem as shown in figure 3.1. We can clearly see the effect of the norm matrix form as a circle boundary, where the behavior of the system changes drastically.

Let us say we wish to know the ground state energy of a system of the form above where the energy gap is equal to roughly 3 and the system is bound. We will say our prior distribution is uniform in both x and y and ranges from -1.5 to 1.5. Our first weight criteria will be that the norm matrix is positive definite, and we set the weight of all samples which violate this condition to 0.

We can now implement additional weight criteria based on our physics considerations. First, we can use a Gaussian term to require that the energy gap be roughly 3

$$w_1 = e^{\frac{\lambda_1 - \lambda_2}{2\sigma}} \tag{3.9}$$
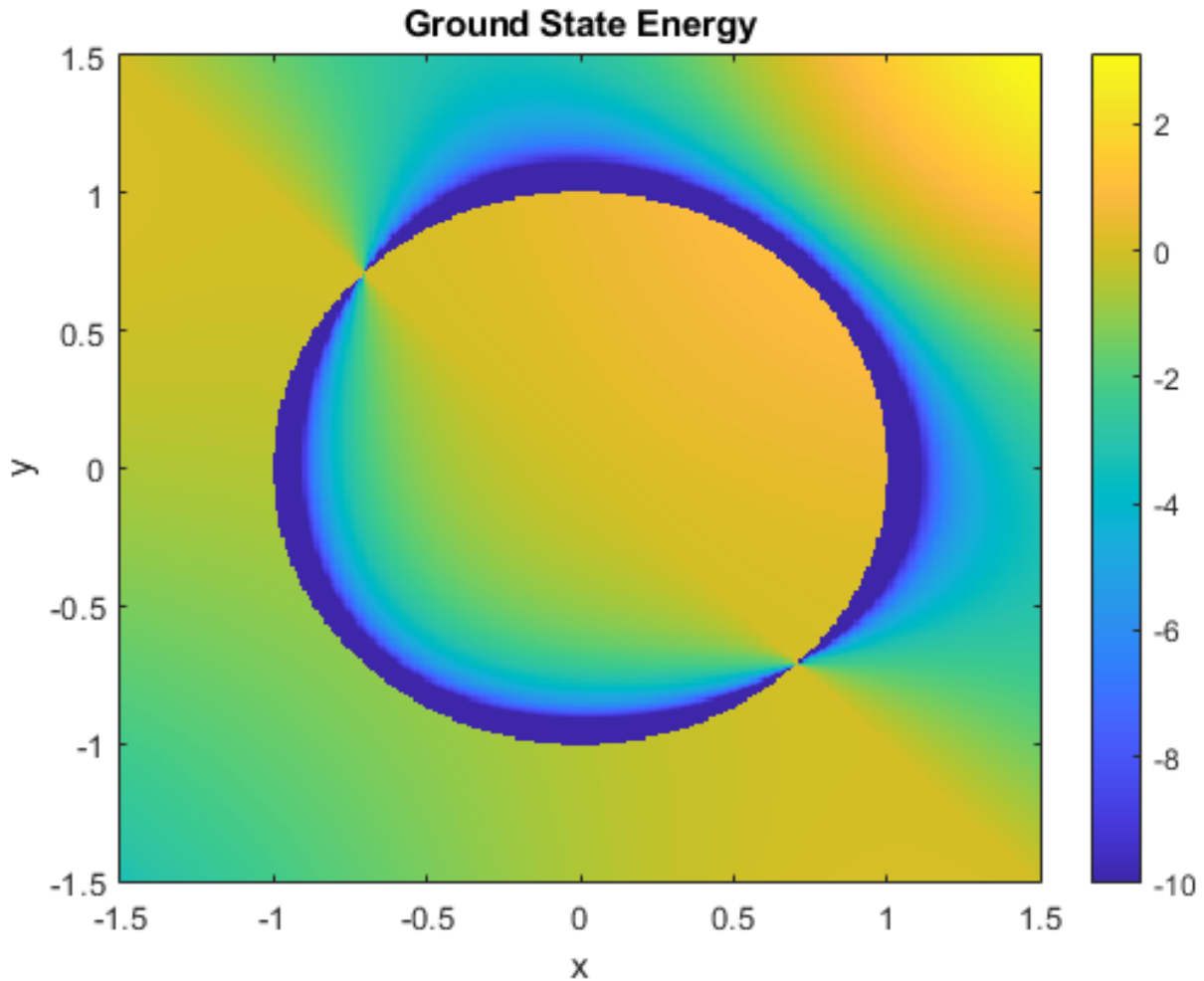
Figure 3.1: Ground state energy in the two parameter generalized eigenvalue problem. The x-axis corresponds to the value of x, and the y-axis corresponds to the value of y. Color represents the value of the ground state energy for values of x and y which correspond with its position on the plot. The clear circle visible corresponds to $\sqrt{x^2 + y^2} = 1$. All systems within that circle have a norm matrix which is positive definite, and systems outside that circle are not positive definite.
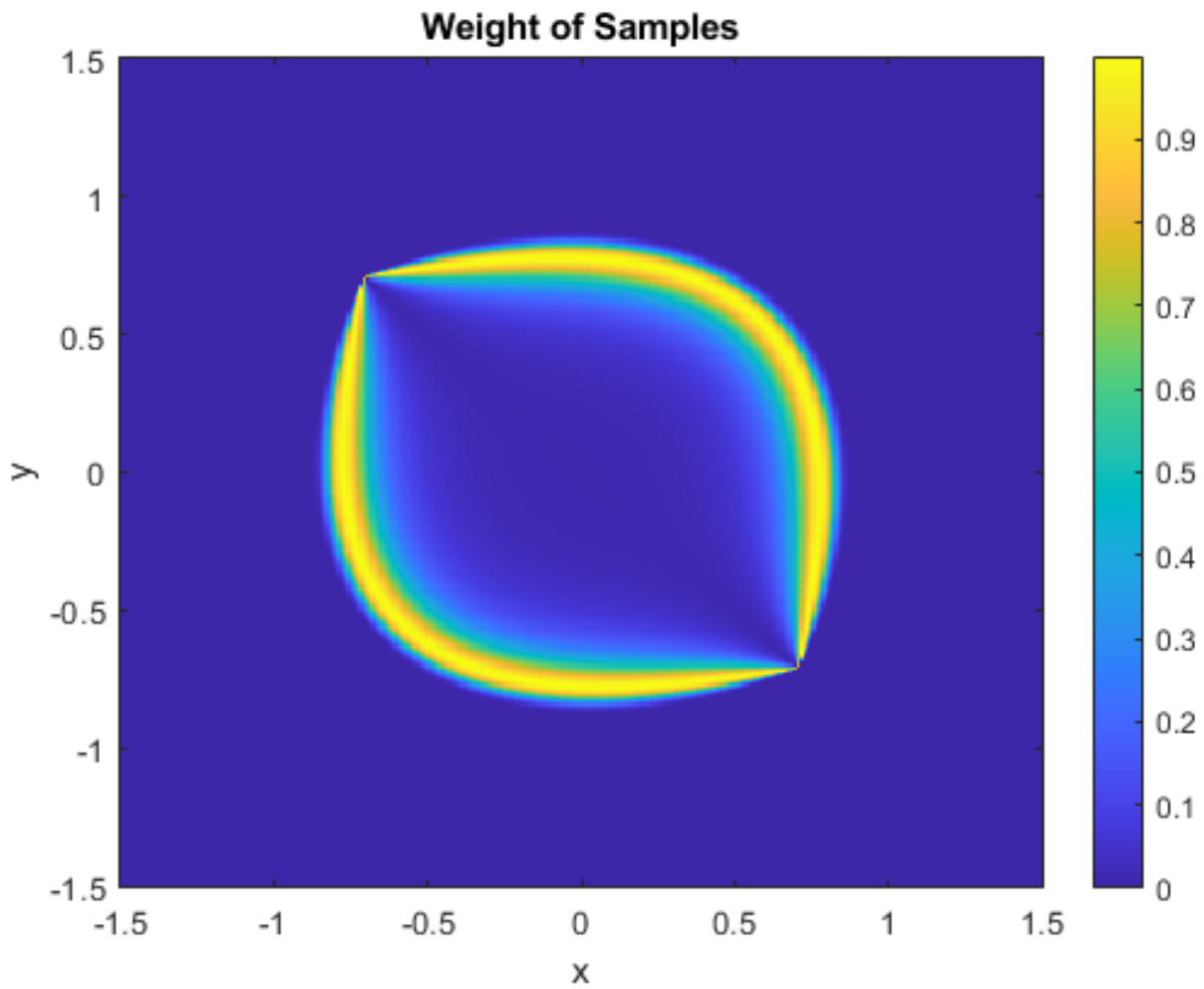
Figure 3.2: Weight for each system given the provided conditions. The x-axis corresponds to the value of x, and the y-axis corresponds to the value of y. The color of each pixel corresponds to the weight of the system at those coordinates. The two clear bands represent two different behaviors this system can take, given our conditions.

Where λ_k is the *kth* energy and σ is a chosen width parameter. These two parameters constrain the possible systems to two small regions, shown in figure 3.2. Each of these regions consists of a thin band, and the behavior of the ground state is quite different in these two states. For the lower left band, the system is bound, while the upper right band corresponds to a continuum state.

To gain further information on these two behaviors, we use a modified variant of DBscan clustering to group final observations into clusters. Note that we are clustering systems by the value of the parameters, x and y, and instead look solely at the ground state energy. This clustering reveals the two regions very clearly in figure 3.3. The blue region represents the bound state, and the energies it can take are a part of a much wider distribution than the orange cluster, which shows a continuum state. We can finally take the weighted median of each cluster, to see that the bound state has a median weight of -1.4118 and the continuum state has a median weight of 0.057. While one might expect the bound state to have a lower energy and the continuum state to have a higher energy based on observation of the plot, the density of points and their associated weight shifts the median.

While this model is extremely basic, it shows the ability of trimmed sampling to find physical, interesting results from a given set of parameters and constraints. In future examples, the degrees of freedom in the system are taken to be each matrix element in the norm and hamiltonian matrices, and so it is entirely impossible to exhaustively scan all solutions in the way of this example. Instead, we make use of sampling techniques, with samples taken according to a prior weight given by the particulars of the problem.

3.5 Bose-Hubbard Model and Eigenvector Continuation

Eigenvector continuation is a powerful and flexible method which is able to reduce a large basis, otherwise intractable by common linear algebra approaches, such as the Lanczos algorithm, to a tractable generalized eigenvalue problem in substantially fewer dimensions. In this example, we apply Eigenvector continuation to the Bose-Hubbard model, introduce errors intentionally to the computation, then correct these errors using trimmed sampling.

3.5.1 Bose-Hubbard Model

For this example, we will consider the Bose-Hubbard model in three-dimensions. This system describes a system of identical bosons on a three-dimensinal lattice. The Hamiltonian contains a hopping coefficient t, an contact interaction coefficient U, and a chemical
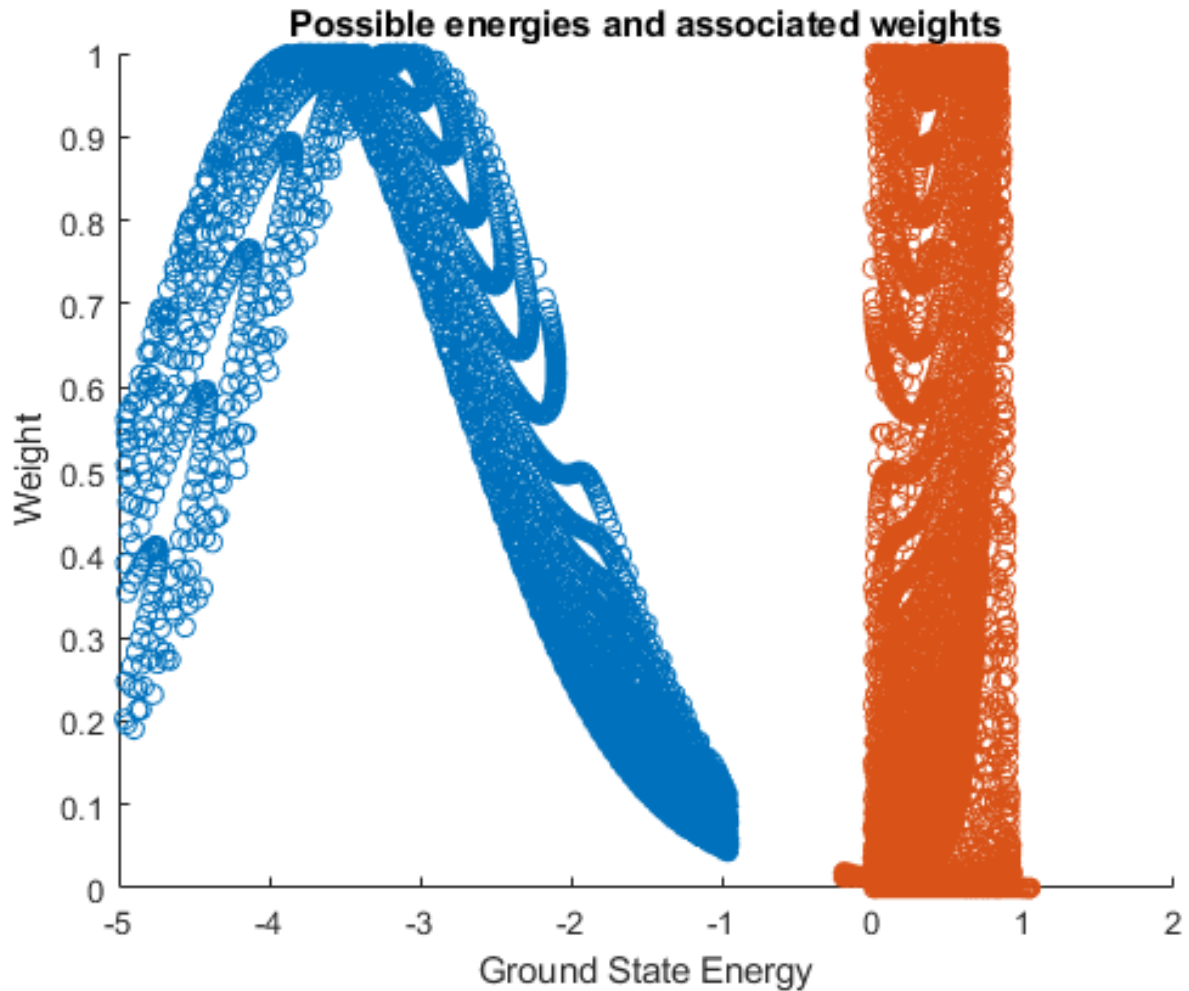
Figure 3.3: Clustering of observations for all samples taken in the 2d toy model generalized eigenvalue problem. The horizontal axis shows the ground state energy for each datapoint, and the vertical axis shows the relative weight of those samples, with the highest weight normalized to 1. Two clusters were found using DBscan. The blue points show one cluster and the orange points show another. A handful of points are removed as a part of the DBscan clustering, as they were not close enough or weighted highly enough to be considered a part of either cluster.

Figure 3.4: Ground state energy of the Bose-Hubbard model as a function of coupling strength U/t. The "exact" ground state energies are plotted as solid lines. The "noiseless EC" data are presented with dashed lines. The "noisy EC" results corresponding with matrix elements $\tilde{H}$ and $\tilde{N}$ are plotted with open circles. The results using "ridge regression" are shown with times symbols. The "raw data" obtained by sampling the prior probability distribution are displayed with open triangles and error bars. The "trimmed sampling" results are plotted as filled circles with error bars.

potential coefficient μ. The Hamiltonian has the form

$$H = -t \sum_{\langle \mathbf{n}', \mathbf{n} \rangle} a^{\dagger}(\mathbf{n}')a(\mathbf{n}) + \frac{U}{2} \sum_{\mathbf{n}} \rho(\mathbf{n})[\rho(\mathbf{n}) - \mathbf{1}] - \mu \sum_{\mathbf{n}} \rho(\mathbf{n}), \qquad (3.10)$$

where $a(n)$ and $a^{\dagger}(n)$ are the annihilation and creation operators for bosons on site n. The first sum runs over all pairs of nearest-neighbors, $\langle n', n \rangle$, and $\rho(n)$ is the density operator defined as $a^{\dagger}(n)a(n)$. For convenience, we set $\mu = -6t$ on a $4 \times 4 \times 4$ lattice, so that the ground state energy of the non-interacting case where $U = 0$ is equal to 0.

3.5.2 Eigenvector Continuation

The chosen method to calculate the ground state energy of this system is Eigenvector Continuation (EC). Eigenvector Continuation is a reduced basis method in which a Hamiltonian controlled by one or more smoothly varying parameters is projected onto the basis of low-lying states of the Hamiltonian in regions of the control parameter which are known. As the control parameters are smoothly varied, the low lying states of the Hamiltonian trace out limited trajectories in finite dimensions, allowing for efficient dimensionality reduction.

For more details, see the subsection in chapter 2.

3.5.3 Results

In this example, we look at the generalized eigenvalue problem associated with Eigenvector Continuation on the Bose-Hubbard Model. In this case, we consider the ground state

eigenvectors for five training values, $U/t = -2.5, -1.9, -1.8, -1.7, -1.6$. These values represent known eigenvectors for the full Hamiltonian, calculated using a more computation intensive method. We then project this Hamiltonian onto these states, resulting in a generalized eigenvalue problem with rank 5. To introduce noise, we then round the entries of each matrix to six decimal places and use these rounded values as the starting matrices, $\tilde{H}$ and $\tilde{N}$. To show that the exact form of the error is unnecessary, rather than a uniform error estimate which would be produced by rounding, we instead assume a Gaussian error distribution with standard deviations $\Delta\tilde{H} = \Delta\tilde{N} = \frac{1}{\sqrt{12}} \times 10^{-6}$.

While eigenvector continuation produces accurate results for this problem without the introduction of rounding or noise, after rounding the matrix elements we see that the character of the ground state, especially for $U/t < 3.5$, is lost, even though matrix elements were only rounded to the sixth digit. This is because the norm matrix becomes increasingly ill-conditioned as the number of basis states increases, and even at only 5 basis states this rounding is sufficient to ruin the results.

We now apply Trimmed Sampling to these results. For the likelihood function, we use the form introduced in 3.4, where the likelihood function is composed of a term ensuring the norm matrix is positive definite, and a term ensuring that the ground state converges as the number of basis states increases. For our estimate of C, the upper bound of the convergence ratio for an exact system, we use $C = 2.5$, though results are not sensitive to this value.

We see in figure 3.4 that Trimmed Sampling does a good job of correcting the error introduced by rounding, and performs better than Ridge Regression. Trimmed Sampling results are shown as filled circles with error bars. We locate the plot symbol at the weighted median and show error bars with lower and upper limits corresponding with the 16th and 84th percentiles respectively. This representation of the distribution is useful, as the distributions have much heavier tails than Gaussian distributions. For these results, we produced 500 samples with non-negligible weights and these small-matrix calculations may be performed easily on a single processor.

For comparison, we also plotted the unweighted results in open triangle symbols. This data represents the effect of sampling the space but calculating the weighted median and associated error bars before applying weighting. In this case, the results are very poor. This indicates that the likelihood function is effectively weighting samples so that those with greater weight produce results closer to the exact results.

Finally, as an additional comparison, we have plotted the results of simple Ridge Regression in $\times$ symbols. Ridge Regression refers to stabilizing the ill-conditioned norm matrix by adding a constant multiple of the identity, ϵI, to the norm matrix. This can significantly reduce problems arising from the unstable inversion of the norm matrix, but introduces an unknown systematic bias. Additionally, the optimal value for ϵ can be diffictult to determine. For these results, we tuned ϵ by hand in an attempt to produce

the best results. We see that even with this hand tuning, Ridge Regression is unable to fully correct the errors.

In these results, we have shown that Trimmed Sampling is an effective method at correcting errors arising from round-off error. With simple filtering criteria which are not sensitive to user choice, we are able to outperform the error correcting abilities of Ridge Regression and reconstruct the phase transition which is lost before any error stabilization is applied. These results do show a systematic underestimation in the region of $U/t < -3.8$, near the avoided level crossing. However, the error estimates give a good estimate of the actual deviation from the exact results.

3.6 Euclidean Time Projection

For another example, we will consider the one dimensional Heisenberg chain. This model emulates the spin interactions of several particles arranged in a chain with periodic boundary conditions.

3.6.1 Heisenberg Model

For this example we will use a one-dimensional Heisenberg Chain with 10 sites The Heisenberg chain is a simple model for a ferromagnetic system and was first introduced in (39). The Hamiltonian for this system is

$$H_J = -J \sum_{j=1}^{N} [\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + \sigma_j^z \sigma_{j+1}^z]. \tag{3.11}$$

Here σ^x, σ^y, and σ^z are the Pauli matrices, N is the number of sites, in this case 10, and J is the coupling to a magnetic field in the z direction. We will calculate the lowest four energy eigenvalues of the subspace where $\sum_j \sigma_j^z = 0$.

3.6.2 Euclidean Time Projection

Euclidean time projection is a method where we start with an initial state, $|v_0\rangle$. We will operate on it with the Euclidean time projection operator, $e^{-H\tau}$ where H is the system Hamiltonian and t is the Euclidean time. We then calculate each new basis vector $|v_n\rangle = e^{-Ht_n}|v_0\rangle$, and project the Hamiltonian onto this basis. For this example, our initial state is an alternating tensor product state, $|\psi\rangle = |0101010101\rangle$. We are using the standard qubit notation where $|0\rangle$ is the $+1$ eigenstate of σ_z and $|1\rangle$ is the -1 eigenstate of σ_z. We will therefore be finding the lowest energy states which have non-zero overlap with this tensor product state. As the state is highly symmetric, we expect this state to have non-zero overlap with a relatively small number of eigenstates.
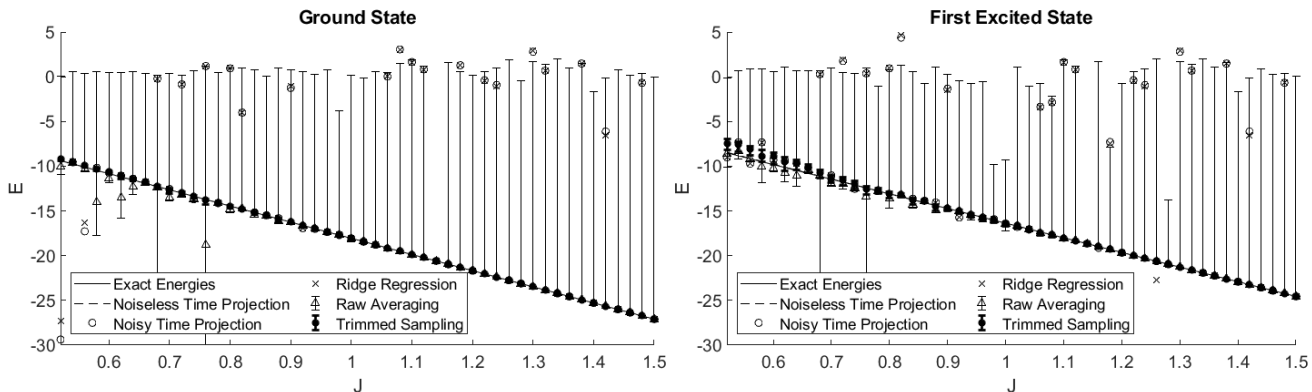
Figure 3.5: Ground state and first excited state energies of the one-dimensional Heisenberg chain as a function of coupling strength J. The "exact" energies are plotted as solid lines. The "noiseless time projection" data are dashed lines. The "noisy time projection" results corresponding with matrix elements $\tilde{H}$ and $\tilde{N}$ are plotted with open circles. The data obtained using "ridge regression" are shown with times symbols. The "raw data" obtained by sampling the prior probability distribution are drawn with open triangles and error bars. The "trimmed sampling" results are plotted as filled circles with error bars.

We operate on $|v_0\rangle$ with the Euclidean time operator $e^-H\tau$ as described in subsection 2.2.2.2. We consider a range of values of J ranging from 0.5 to 1.5. For each value of J, we take a series of Euclidean time slices $\tau_n = 0.1n$ for $n = 0, 1, 2, 3, 4, 5$. We define each basis vector as the Euclidean time projection to those time slices, $|v_n\rangle = e^{-H\tau_n}|v_0\rangle$. We calculate the projected Hamiltonian and Norm matrix and solve the resulting generalized eigenvalue problem.

3.6.3 Results

To introduce noise, we apply random Gaussian noise with standard deviation $\sigma = 0.01$ to each element of both the Hamiltonian and the Norm matrices. The noisy norm and H matrices are our initial estimates, $\tilde{N}$ and $\tilde{H}$, and assume the uncertainty estimates $\Delta\tilde{H} = \Delta\tilde{N} = \sigma = 0.01$. We then solve the generalized eigenvalue problem to calculate the lowest four energy eigenvalues as well as some observables, in this case the spin operators $\langle\sigma_1^z\sigma_2^z\rangle$ and $\langle\sigma_1^z\sigma_3^z\rangle$.

Because J is an overall scale factor, the Euclidean time calculations are different for each J, but the energy eigenvalues are linear in J. The results of this calculation is presented in 3.5. The "exact" energies plotted with solid lines are calculated using direct diagonalization. The "noiseless time projection" results are shown with dashed lines and show the results of solving the generalized eigenvalue problem before any noise is applied. The "noisy time projection" results show the results of solving the generalized eigenvalue problem after gaussian noise is applied to each element and are plotted with open circles. The results for "ridge regression" are shown with × symbols. These results are obtained
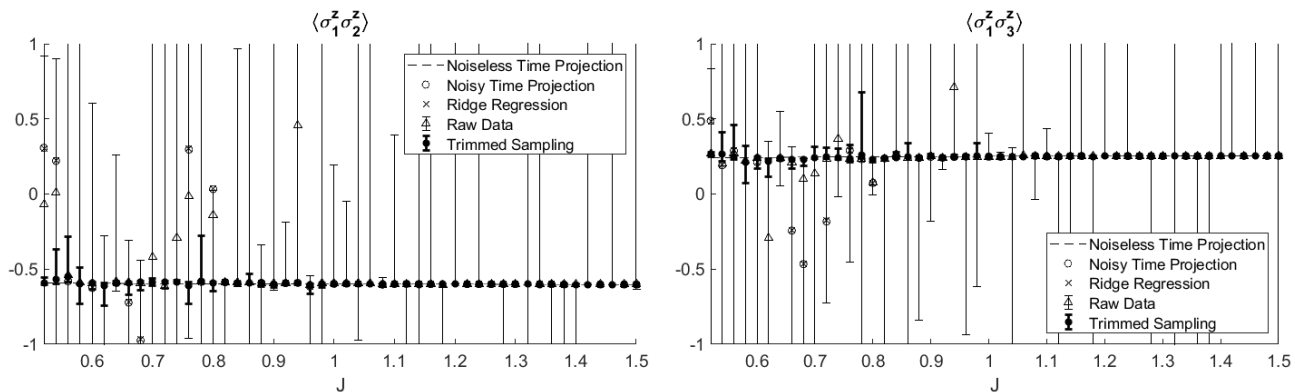
Figure 3.6: Spin pair expectation values for the ground state of the Heisenberg model as a function of J. $\langle\sigma_1^z\sigma_2^z\rangle$ is shown in the left panel, and $\langle\sigma_1^z\sigma_3^z\rangle$ is presented in the right panel. The plot symbols are the same as in Fig. 3.5.

by hand tuning the nugget ϵI to produce the best results, but these results showed only marginal improvement over the noisy case. The "raw data" is shown in open triangles and is obtained by directly sampling the prior probability distribution associated with the mean values of $\tilde{H}$ and $\tilde{N}$.

The "trimmed sampling" results are shown as filled circles with dark error bars. These results sample the prior distribution, similar to the "raw data" results, but apply the weighting criteria as described in 3.4. We have used the value of $C = 2.5$ for $f_c(H, N)$, exactly as in the Bose-Hubbard example. Trimmed sampling substantially reduces the error associated with the added noise, and provides error bars that give a reasonable estimate of the actual deviation from the noiseless results. In contrast, ridge regression does not give consistently reliable results for this benchmark test.

We also see these improved results when considering the first excited state, even though the filtering criteria only considers the ground state. This indicates that trimmed sampling is correctly reducing the errors associated with noise for the full matrix.

In addition to calculating the energies, we can also compute spin observables for this system. In Fig 3.6 we show the ground-state expectation value of the product of nearest neighbor spins, $\langle\sigma_1^z\sigma_2^z\rangle$ in the left panel, and next-to-nearest neighbors, $\langle\sigma_1^z\sigma_3^z\rangle$ in the right panel. Again we see good improvement of the error compared to ridge regression and the noisy results, and the error estimates give a good estimate of the real deviation from exact results.

In this model as well we see that trimmed sampling is an effective method to reduce the effect of noise on the Heisenberg system. In addition to calculating the ground state energy, we are able to use the same corrected matrices to calculate excited states and expectation values, indicating the Gausssian noise introduced has been reduced substantially.

3.7 Improving Sample Performance

While basic sampling is sufficient for many applications, in some cases it is desirable to improve the performance for use in larger or more difficult situations. This sections describes some methods to improve the performance of the sampling.

3.7.1 Markov-Chain Monte Carlo

Results in the previous sections have used a simple stochastic sampling method, where each matrix element is independently sampled according to its associated error estimate. Stochastic sampling is sufficient when the dimension of the matrices is small (typically less than or close to 10), but when the dimension of the matrices grows the space becomes too large to effectively sample this way.

In further examples, we will make use of Markov Chain Monte Carlo to sample the posterior distributions. We begin with the initial matrix estimates, $\tilde{H}$ and $\tilde{N}$. We then take a small permutation of each matrix element, similar to stochastic sampling, but in this case the prior distribution we draw from is much smaller than the error estimates. These new matrices are called $\tilde{H}_{trial}$ and $\tilde{N}_{trial}$.

We calculate likelihood functions of these matrices, but also include a term $f_{prior}(H, N)$ which equals the likelihood of the sample given only the prior error estimates on the elements of $\tilde{H}$ and $\tilde{N}$. This term takes the form

$$f_{prior}(H, N) = \frac{1}{2\pi\Delta\tilde{H}\Delta\tilde{N}} \exp\left(-\frac{1}{2}\sum_{i,j} \frac{(\tilde{H}_{i,j} - \tilde{H}_{trial,i,j})^2}{\Delta\tilde{H}} + \frac{(\tilde{N}_{i,j} - \tilde{N}_{trial,i,j})^2}{\Delta\tilde{N}}\right) \quad (3.12)$$

We then use the Metropolis-Hastings algorithm to accept or reject each sample. We compare the likelihood functions of the current sample and the previous accepted sample and accept the trial sample with probability $P = \min(e^{(w_{trial}-w_{prev})/T}, 1)$ where w_{prev} is the previous sample's likelihood, w_{trial} is the trial sample's likelihood, and T is some non-zero temperature. If a sample is accepted, subsequent samples are drawn from a small shift from that sample, rather than the initial matrices $\tilde{H}$ and $\tilde{N}$. This allows the chain to explore the space and find regions of high likelihood.

3.7.2 Nugget Regularization Start

In highly singular systems when using Markov-Chain Monte Carlo sampling, locating a sample where the norm matrix is positive definite can take many iterations. However, once such a sample is found, it is often very fast to sample the surrounding region and find samples of good likelihood. To accelerate the initial sampling, we add a small multiple of the identity, ϵI, to the norm matrix prior to beginning sampling and use this regularized norm matrix as the first $\tilde{N}_{trial}$.

While nugget regularization can introduce unknown biases and be difficult to properly tune, this is not a problem in this case because it is only an initial guess in the Markov Chain. By starting the chain at this regularized norm matrix, we ensure that the chain will quickly find a region of positive definite norm matrices and can then optimally sample non-negligible trial matrices.

3.8 Improving Filtering Criteria

This section will explain how we can improve the filtering criteria and what kinds of criteria are useful. This may be a good section to include the non-bifurcated Lipkin model, and how enforcing the symmetry via the criteria gives good results.

3.9 Lipkin Model and Generator Coordinate Method

As a final example, we will look at the Lipkin Model, a toy model used to emulate some of the many-body dynamics near the Fermi surface. We will use the Generator Coordinate Method to reduce the model to a tractable size and examine the effects of several filtering criteria on the final results.

3.9.1 Lipkin Model

The Lipkin Model describes a set of N fermions in two N-fold degenerate levels separated by an energy ϵ. Each state is labeled by the quantum numbers $\sigma = \pm 1$, to denote whether the particle is in the upper or lower level, and m to specify the particular degenerate state within that level. The Lipkin Model Hamiltonian is:

$$H = \frac{1}{2}\epsilon \sum_{m=1}^{\Omega} \sum_{\sigma=-1}^{+1} \sigma a_{m\sigma}^{+} a_{m\sigma} - \frac{1}{2}V \sum_{mm'=1}^{\Omega} \sum_{\sigma=-1}^{+1} a_{m\sigma}^{+} a_{m'\sigma}^{+} a_{m'-\sigma} a_{m-\sigma} \tag{3.13}$$

where $a_{m\sigma}^{+}$ creates a particle in the σ level at position m, V is the interaction strength which moves a pair of particles between levels, and Ω is the number of degenerate states in each level.

The Lipkin Model has the benefit that it is exactly solvable. Because of this, it provides an excellent benchmark against which to text numerical methods. To compute the exact solution, we will use the symmetry method introduced earlier. We introduce quasispin operators K_0, K_+, and K_-.

K_0 represents the difference between the number of particles in the upper level and the number of particles in the lower level. It is written as

$$\hat{K}_0 = \frac{1}{2} \sum_{m=1}^{\Omega} (a_{m+}^{+} a_{m+} - a_{m-}^{+} a_{m-}) \tag{3.14}$$

Figure 3.7: A rough schematic of the Lipkin model showing the energy levels and labeling of states.

K_+ is the operator which increases the value of the quasispin operator K_0 and is written as

$$\hat{K}_+ = \sum_{m=1}^{\Omega} a_{m+}^+ a_{m-} \tag{3.15}$$

and K_- is the hermitian conjugate of K_+, $\hat{K}_- = (\hat{K}_+)^\dagger$. With these operators, we can express the Lipkin Model Hamiltonian as

$$\hat{H} = \epsilon \hat{K}_0 - \frac{1}{2} V (\hat{K}_+ \hat{K}_+ - \hat{K}_- \hat{K}_-) \tag{3.16}$$

These quasispin operators follow the usual angular momentum commutation relations, $[\hat{K}_+, \hat{K}_-] = 2\hat{K}_0$ and $[\hat{K}_0, \hat{K}_\pm] = \pm \hat{K}_\pm$. By expressing the Hamiltonian in a basis of eigenstates of $\hat{K}_0$, we reduce the dimensionality of the problem significantly and allow it to be solved by diagonalizing the Hamiltonian.

We will consider the half filled case, where $n = \Omega$. Because the particle number is fixed, the maximum quasispin is $k = \Omega/2$. This corresponds to the state where all particles are in the upper level. The Hamiltonian cannot mix states with different maximum quasispin, so we can construct a basis out of only the states which have equal maximum quasispin. We will label the basis states $|kk_z\rangle$ where $k_z = -k, -k+1, \cdots, k-1, k$. There are thus $2\Omega + 1$ total basis states in the half filled case we are considering.

We will make use of standard angular momentum relations, namely

$$\hat{K}_0 |kk_z\rangle = k_z |kk_z\rangle \tag{3.17}$$

and

$$\hat{K}_\pm |kk_z\rangle = \sqrt{k(k+1) - k_z(k_z \pm 1)} |kk_z + 1\rangle \tag{3.18}$$

We can now compute the matrix elements for the Hamiltonian. These quasispin states are orthogonal to each other, so

$$\langle kk_z | \epsilon \hat{K}_0 | kk_z' \rangle = \delta_{k_z k_z'} \epsilon k_z \tag{3.19}$$
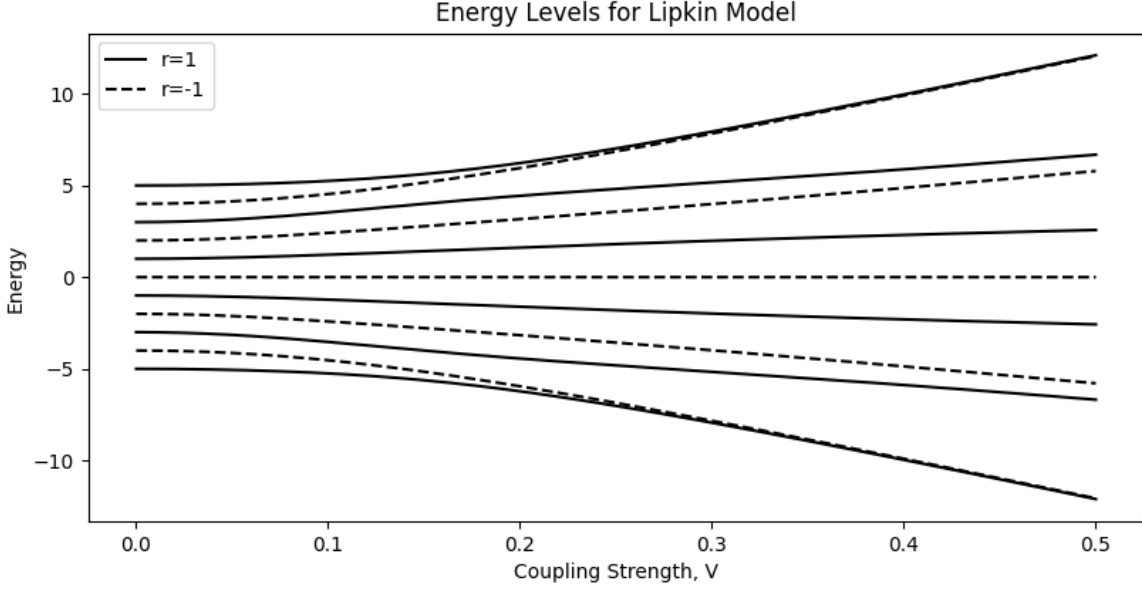
Figure 3.8: The exact energy levels of the Lipkin model with respect to a varying copuling strength, V. We show all energy levels for a half filled system with $\Omega = 10$, $\epsilon = 1$, and $n = \Omega$. The solid lines show the solution with positive signature, and the dashed lines show the solutions with negative signature.

The $\hat{K}_0$ operator does not change quasispin, so it is diagonal in this basis. Since the interaction term always changes the quasispin by ± 2, it must be zero everywhere except with respect to two states which differ by exactly 2 quasispin. We have

$$\langle k k_z | \hat{K}_+ \hat{K}_+ | k k_z' \rangle = -\frac{V}{2} \delta_{k_z+2,k_z'} \sqrt{k(k+1) - k_z'(k_z'+1)} \sqrt{k(k+1) - (k_z'+1)(k_z'+2)} \tag{3.20}$$

and similarly for the lowering operators,

$$\langle k k_z | \hat{K}_- \hat{K}_- | k k_z' \rangle = -\frac{V}{2} \delta_{k_z-2,k_z'} \sqrt{k(k+1) - k_z'(k_z'-1)} \sqrt{k(k+1) - (k_z'-1)(k_z'-2)} \tag{3.21}$$

An additional symmetry we can take advantage of is signature. The signature operator

$$R = e^{i\pi \hat{K}_0} \tag{3.22}$$

commutes with the Hamiltonian. This operator has, for the half filled, even particle number case, only two possible eigenvalues, $r = \pm 1$. We see that when k_z is even, signature is positive, and when k_z is odd signature is negative. This allows us to further divide the system into two sections with positive and negative signature. More details are explained later when we use this signature symmetry and the generator coordinate method.

It is thus quite simple to construct the matrix in this representation. Figure 3.8 shows the results of this direct solution. We show both the positive and negative signature

results for $k_z = \Omega/2 = 5$. We see that at very low values of the coupling, the Hamiltonian simply counts the difference between particles on the upper and lower levels. As the paring strength increases, the energies split and take on more complicated behavior. At very strong values of the coupling, positive and negative signature states begin to pair up and their energies become increasingly degenerate.

3.9.2 Generator Coordinate Method

The Generator Coordinate Method is a method to reduce a large basis to a reduced basis of known states. We construct a linear superposition of many product wave functions. The trial wavefunction is a superposition of the generating functions $|\tau\rangle$ which may be functions of multiple parameters. The ansatz is an integral over these parameters

$$|\Psi\rangle = \int d\tau f(\tau)|\tau\rangle \tag{3.23}$$

where $f(\tau)$ are weight functions.

For this example, we will be using coherent states as the basis. We define these basis states as

$$|\tau\rangle = \frac{1}{(1 + |\tau|^2)^j} e^{\tau \hat{K}_+} |k - k\rangle \tag{3.24}$$

We will be considering the half filled case, where $j = \Omega/2$. The parameter τ is defined as $\tau = \tan\frac{\theta}{2}$. The angle θ is the parameter for these basis states.

The expectation value of the Hamiltonian with these trial wavefunctions is

$$E = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int d\tau d\tau' f(\tau) f(\tau') \langle\tau|\hat{H}|\tau'\rangle}{\int d\tau d\tau' f(\tau) f(\tau') \langle\tau|\tau'\rangle} \tag{3.25}$$

We define $H(\tau, \tau') = \langle\tau|\hat{H}|\tau'\rangle$ and $N(\tau, \tau') = \langle\tau|\tau'\rangle$. Minimizing this energy with respect to the parameters $f(\tau)$, we get

$$\frac{\partial E}{\partial f(\tau)} = 0 = \frac{\int d\tau' H(\tau, \tau') f(\tau) - E \int d\tau' N(\tau, \tau') f(\tau')}{\int d\tau d\tau' f(\tau) f(\tau') N(\tau, \tau')} \tag{3.26}$$

Rearranged, this gives the Hill-Wheeler equations

$$\int d\tau' H(\tau, \tau') f(\tau) = E \int d\tau' N(\tau, \tau') f(\tau) \tag{3.27}$$

By discretizing the parameter θ, and therefore τ, we can formulate a matrix representation of these equations in the form of a generalized eigenvalue problem. To do so, we need to calculate the matrix elements for each operator $\langle\tau|\hat{K}_0|\tau'\rangle$, $\langle\tau|\hat{K}_+^2|\tau'\rangle$, and the normalization, $\langle\tau|\tau'\rangle$.

To begin, we need to express $|\tau'\rangle$ in terms of $|\tau\rangle$. As $|\tau\rangle = \frac{1}{(1+|\tau'|^2)^k} e^{\tau'\hat{K}_+}|k - k\rangle$, we can express $|\tau'\rangle$ as

$$\frac{(1 + |\tau|^2)^k}{(1 + |\tau'|)^k} e^{(\tau' - \tau)\hat{K}_+} \frac{1}{(1 + |\tau|^2)^k} e^{\tau\hat{K}_+}|k - k\rangle = \frac{(1 + |\tau|^2)^k}{(1 + |\tau'|^2)^k} e^{(\tau' - \tau)\hat{K}_+}|\tau\rangle \tag{3.28}$$

Now to calculate the matrix elements, we will make use of generating functions and a result from the Thouless Theorm,

$$\langle\tau|e^{\alpha_- J_-}e^{\alpha_0 J_0}e^{\alpha_+ J_+}|\tau\rangle = (1+|\tau|^2)^{-2j}[e^{-\frac{1}{2}\alpha_0} + e^{\frac{1}{2}\alpha_0}(\tau^\dagger + \alpha_-)(\tau + \alpha_+)]^{2j} \tag{3.29}$$

.

First, we can see that

$$N(\tau,\tau') = \langle\tau|\tau'\rangle = \langle\tau|\frac{(1+|\tau|^2)^k}{(1+|\tau'|^2)^k}e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle \tag{3.30}$$

Applying Thoules theorem and $k = \Omega/2$, because we are considering the half-filled case, gives

$$N(\tau,\tau') = \frac{(1+|\tau|^2)^{-\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}(1+\tau^\dagger\tau')^\Omega. \tag{3.31}$$

Because $\tau = \tan\theta/2$, we an rewrite this as $N(\theta,\theta') = [\cos\frac{\theta-\theta'}{2}]^\Omega$.

To calculate the Hamiltonian kernel, $H(\tau,\tau')$, we need to compute the expectation values $\langle\tau|\hat{K}_0|\tau'\rangle$ and $\langle\tau|\hat{K}_+^2|\tau'\rangle$. Beginning with $\langle\tau|\hat{K}_0|\tau'\rangle$ we see

$$\langle\tau|\hat{K}_0|\tau'\rangle = \frac{(1+|\tau|^2)^{\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}\langle\tau|\hat{K}_0 e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle \tag{3.32}$$

We apply generator functions to rewrite the expectation value as

$$\langle\tau|\hat{K}_0 e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle = \frac{\partial}{\partial\alpha_0}\langle\tau|e^{\hat{K}_0}e^{(\tau'-\tau)\hat{K}_+}|\tau'\rangle|_{\alpha_0=0} = \frac{\partial}{\partial\alpha_0}[(1+|\tau|^2)^{-\Omega}(e^{-\frac{1}{2}\alpha_0}+e^{\frac{1}{2}\alpha_0}\tau^\dagger\tau')^\Omega]_{\alpha_0=0} \tag{3.33}$$

Putting it back together and simplifying gives

$$\langle\tau|\hat{K}_0|\tau'\rangle = -\frac{(1+|\tau|^2)^{-\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}\frac{\Omega}{2}(1-\tau\tau')(1+\tau\tau')^{\Omega-1} \tag{3.34}$$

Applying a similar procedure to find $\langle\tau|\hat{K}_+^2|\tau'\rangle$, we again use generator functions

$$\langle\tau|\hat{K}_+^2|\tau'\rangle = \frac{(1+|\tau|^2)^{\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}\langle\tau|\hat{K}_+^2 e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle \tag{3.35}$$

Again looking just at the expectation value,

$$\langle\tau|\hat{K}_+^2 e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle = \frac{\partial^2}{\partial\alpha_+^2}\langle\tau|e^{\alpha_+ \hat{K}_+}e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle|_{\alpha_+=0} \tag{3.36}$$

Now applying the result in eq.3.29 we find

$$\langle\tau|\hat{K}_+^2 e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle = \frac{\partial^2}{\partial\alpha_+^2}[(1+|\tau|^2)^{-\Omega}(1+\tau(\alpha_+ + \tau'))^\Omega]_{\alpha_+=0}. \tag{3.37}$$

Simplifying this result and putting it together gives

$$\langle\tau|\hat{K}_+^2|\tau'\rangle = \frac{(1+|\tau|^2)^{-\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}\Omega(\Omega-1)(\tau')^2(1+\tau\tau')^{\Omega-2}. \tag{3.38}$$

We can calculate $\langle\tau|\hat{K}_-^2|\tau'\rangle$ as the conjugate of $\langle\tau|\hat{K}_+|\tau'\rangle$. With these calculations, we can express the expectation value of the Hamiltonian in terms of θ and θ',

$$H(\theta,\theta') = -\frac{\epsilon\Omega}{2}N(\tau,\tau')\left[\frac{\cos\frac{\theta+\theta'}{2}}{\cos\frac{\theta-\theta'}{2}} + \frac{\chi}{2}\left(\frac{1+\sin^2\frac{\theta+\theta'}{2}}{\cos^2\frac{\theta-\theta'}{2}} - 1\right)\right] \tag{3.39}$$

Where we have defined $\chi = \frac{V}{\epsilon}(\Omega-1)$.

For this example, we will also be looking at the $\hat{K}_x$ operator, defined as $\hat{K}_x = \frac{1}{2}(\hat{K}_+ + \hat{K}_-$. We will also look at its square, $\hat{K}_x^2$. Following the procedure above, we begin by finding the expression

$$\langle\tau|\hat{K}_+|\tau'\rangle = \frac{(1+|\tau|^2)^{\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}\langle\tau|\hat{K}_+ e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle \tag{3.40}$$

Making use of Thouless Theorem,

$$\langle\tau|\hat{K}_+|\tau'\rangle = a\frac{\partial}{\partial\alpha_+}\langle\tau|e^{\alpha_+\hat{K}_+}e^{(\tau'-\tau)\hat{K}_+}|\tau\rangle|_{\alpha_+=0} \tag{3.41}$$

Simplifying the result, we find

$$\langle\tau|\hat{K}_+|\tau'\rangle = \frac{(1+|\tau|^2)^{-\Omega/2}}{(1+|\tau'|^2)^{\Omega/2}}\Omega\left[1+\tau'\tau^*\right]^{\Omega-1}\tau^* \tag{3.42}$$

Since we know that $\langle\tau|\hat{K}_+|\tau'\rangle^\dagger = \langle\tau|\hat{K}_-|\tau'\rangle$, we can find

$$\langle\tau|\hat{K}_x|\tau'\rangle = \frac{\Omega}{2}\langle\tau|\tau'\rangle\frac{\tan\frac{\theta'}{2}+\tan\frac{\theta}{2}}{1+\tan\frac{\theta'}{2}\tan\frac{\theta}{2}} \tag{3.43}$$

Naïvely, we can attempt to calculate the ground state energy and the expectation value of these operators of this system by taking discrete samples of θ and θ', calculating the Norm and Hamiltonian matrices from the kernels, and diagonalizing the resulting generalized eigenvalue problem. However, this process has significant stability issues. As we increase the number of basis states by sampling θ and θ' at finer increments, the norm matrix becomes increasingly ill-conditioned, and as a result this direct diagonalization can fail.

Figure 3.9 shows the relative error between the direct Hill Wheeler diagonalization and the exact calculation for the ground state and first excited state energies for $\Omega = 100$, $\chi = 1.2$, and an increasing number of samples. As expected, increasing the number of basis vectors reduces the error until the matrix singularity becomes a significant problem, in this case at approximately 80 samples. This is assuming that each vector taken is evenly spaced and that there are no degeneracies in the basis vectors.

Figure 3.10 shows the relative error between the direct Hill Wheeler diagonalization and the exact calculation for the ground state and first excited state energies, but this time, when selecting basis vectors, a single degenerate vector is added by including both $\theta = -\pi$ and $\theta = \pi$. This causes the instability problem to become rapidly apparent at much smaller basis sizes. We will show one set of results with this added degeneracy, and another set where error is added by truncating matrix elements instead.

Figure 3.9: The logarithm of the error for the ground state and first excited state of the Lipkin Model discretized directly. As the number of values of θ increases, we see that the error of these two states decreases until it reaches machine precision. Further increasing the number of values of θ causes the matrix to become singular at approximately 80 values, in this case. Beyond this, the results become badly error prone.
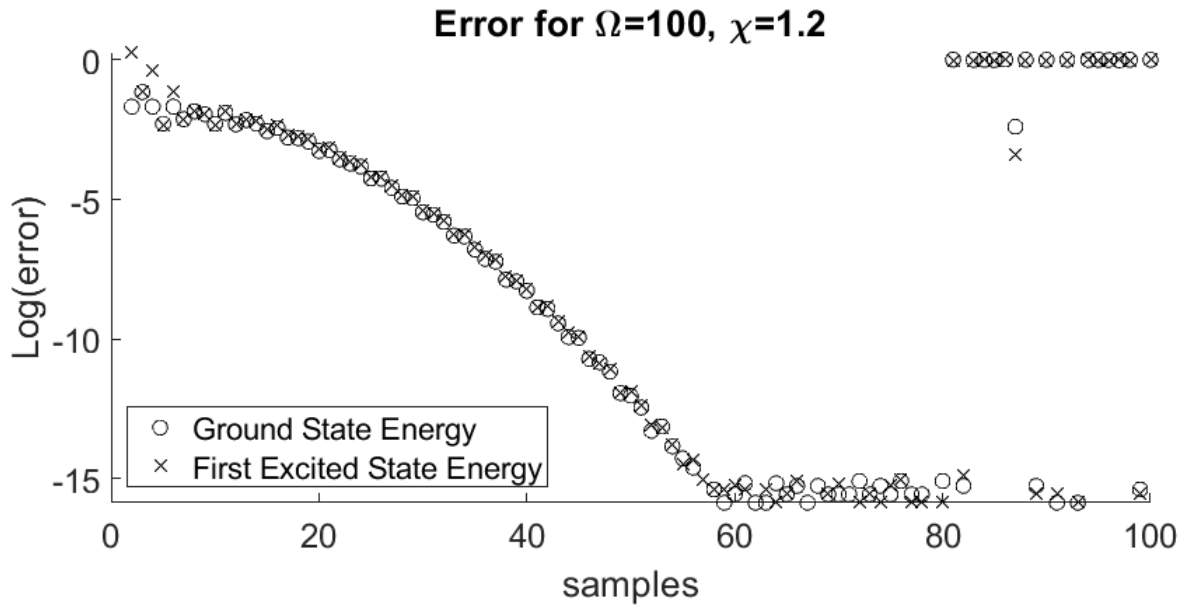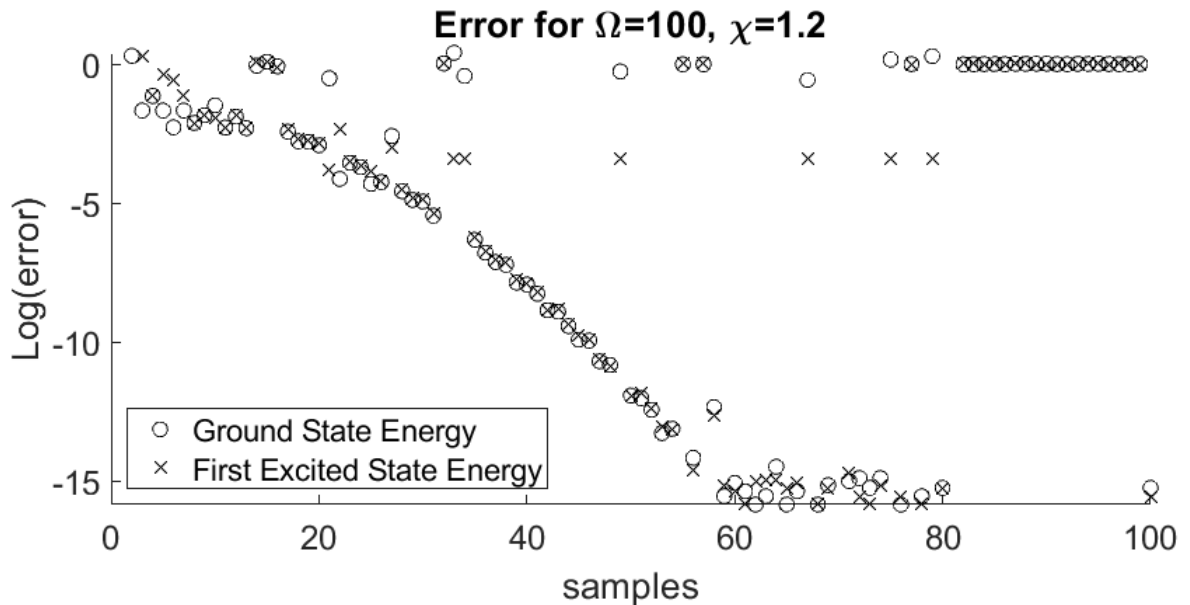

Figure 3.10: The logarithm of the error for the ground state and first excited state of the Lipkin Model discretized directly with an additional degenerate vector added. We see that, similar to the case without a degenerate vector added, the error decreases towards machine precision, but now has many sporadic errors caused by the ill-conditioning.

47

3.9.2.1 Symmetric orthogonalization

A standard approach to address this problem is called symmetric orthogonalization. In this approach, we will attempt to diagonalize the norm matrix, allowing us to calculate a Hamiltonian matrix in the basis of this diagonalized norm matrix and avoid the problems of inverting a nearly singular matrix. To begin, we expand $N(\tau, \tau') = \langle \tau | \tau' \rangle = \cos \frac{\theta - \theta'}{2}$ as a sum of exponentials,

$$\mathcal{N}(\tau, \tau') = \sum_{k=-\Omega/2}^{\Omega/2} e^{-ik\theta} \frac{1}{2^\Omega} \begin{pmatrix} \Omega \\ k + \frac{\Omega}{2} \end{pmatrix} e^{ik\theta'} \tag{3.44}$$

From this we can see that the eigenvalues of N are $n_k = \frac{2\pi}{2^\Omega} \binom{\Omega}{k+\Omega/2}$. We can now calculate the matrix elements for the projected Lipkin Model Hamiltonian with the diagonalized norm matrix:

$$(\hat{H}_{LM})_{kk'} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta d\theta' \frac{e^{ik\theta}}{\sqrt{n_k}} H(\tau, \tau') \frac{e^{-ik'\theta'}}{\sqrt{n_{k'}}} \tag{3.45}$$

This leaves us with a matrix representation of the Hamiltonain in the momentum basis, and we no longer need to invert the norm matrix directly. However, this comes at the cost of now needing to calculate $(\Omega + 1)\Omega/2$ separate integrals, each over the full range of θ values. While this can be done numerically with high precision, it quickly becomes impractical for large systems.

To address the computational scaling, we can instead consider only the k-basis states which contribute most to the overall solution; that is, we can disregard instances where the norm matrix's eigenvalues are very small. This allows us to calculate far fewer integrals overall. We choose to cut off all but the n most significant values of k, and this allows us to directly compare the performance of this method to trimmed sampling. It is worth noting however that while we add noise to the direct discretized matrices, it is not simple to directly add noise to the matrices obtained using symmetric orthogonalization in a way which can be directly compared. However, symmetric orthogonalization becomes impractical at matrix sizes much smaller than those tractable with trimmed sampling.

3.9.3 Degenerate Basis

For the first example of the Lipkin model, we will start with a half-filled system and $\Omega = 200$. For these tests, we will introduce error into the direct Hill-Wheeler calcualtion by including an additional degenerate basis vector by including both $\theta = -\pi$ and $\theta = \pi$.

For posterior likelihood functions, we use a product of three functions. The first two are similar to previous examples. We use the condition that the norm matrix must be positive definite, $w_1 = \Theta(\lambda_{n,0})$. We also use a similar condition to the convergence criteria for the ground state. In this case, rather than the cutoff ratio as is used in previous examples, we instead look at the square of the second finite difference of the
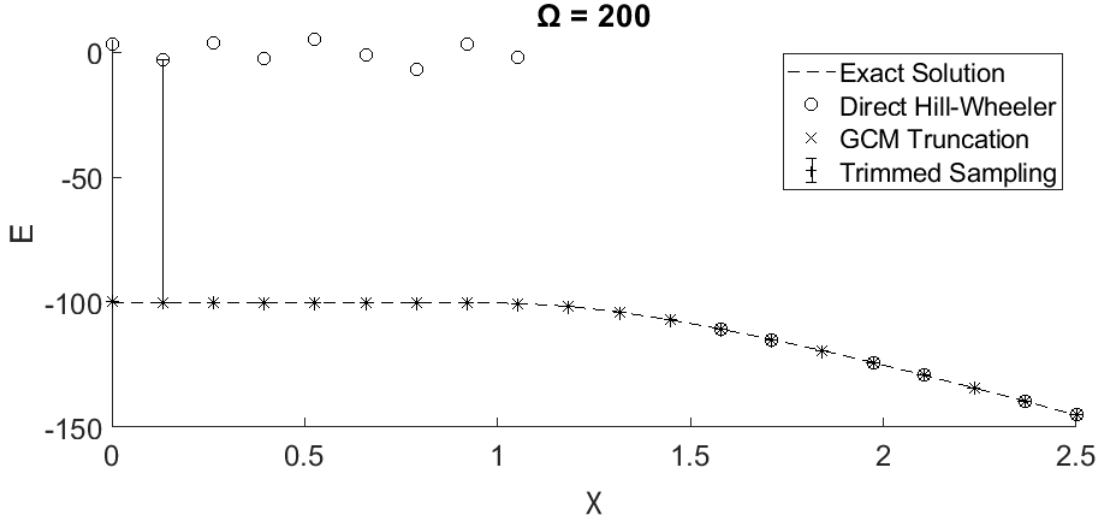
Figure 3.11: The ground state energy for the Lipkin Model. The exact solution is shown in dashed lines. Results from direct discretization are shown in open circles. Results from using k-truncation and norm matrix diagonalization are shown in 'x' symbols. Finally, results for Trimmed Sampling are shown with '+' symbols. The error estimates for Trimmed Sampling are extremely small for most data points. The significantly larger error bar for the second data point indicates that the distribution of accepted points has very wide tails. At a number of points, no value for direct discretization is shown. This is because the norm matrix became so badly singular that no value was able to be calculated at all.

sequence of energies as we reduce the basis size. For this example, we reduce the basis size by removing every $n/2$ samples, then every $n/2 - 1$ samples, etc. This keeps the spacing of basis vectors consistent in θ. We use the second finite difference here to focus on removing any matrix sets where increasing the number of basis vectors causes significant, rapid changes in the ground state energy.

The final condition we use is the expectation value of the $\hat{K}_x$ operator in the ground state. Because the Lipkin model's interaction term is quadratic in raising and lowering operators, the expectation value of the $\hat{K}_x$ operator must be zero for any eigenstate of the Hamiltonian. We can thus use a posterior likelihood which is an exponential in the square of this expectation value to penalize the weight of matrix pairs which fail this condition,

$$w_3 = \exp(-\langle\psi_0|\hat{K}_{x,\text{trial}}|\psi_0\rangle^2)$$

3.9.3.1 Results

Results for the ground state energy are shown in figure 3.11. Trimmed sampling shows significant improvement of the ground state energy over the results from direct diagonalization, especially at lower values of the coupling. Trimmed sampling performs similar to symmetric orthogonalization in this case, but as is shown in figure 3.14, it achieves
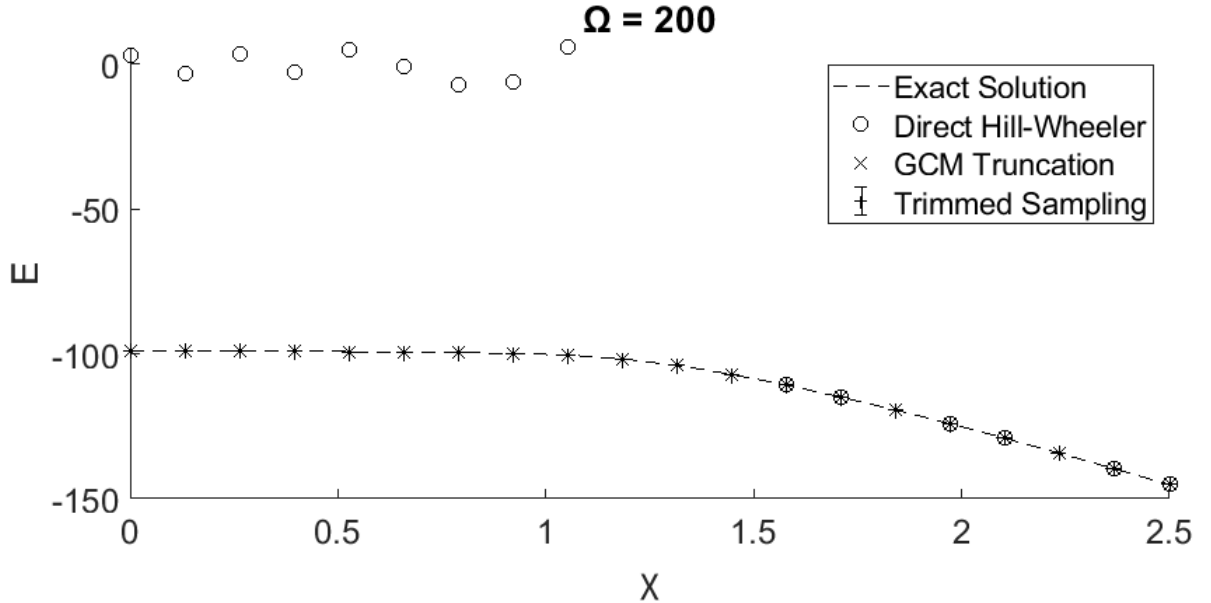
Figure 3.12: First excited state energy for the Lipkin Model. The exact value is shown in dashed lines. Direct θ discretization is shown in open circles. Norm matrix diagonalization with k-truncation is shown in 'x' symbols, and Trimmed Sampling results are shown with '+' symbols. We see that the Trimmed Sampling produces excellent error correction over the Direct θ discretization of the Hill-Wheeler equations.

this accuracy while requiring only a small fraction of the time required for symmetric orthogonalization.

Results for the first excited state energy are shown in figure 3.12. We see similar performance as with the ground state energy, with trimmed sampling outperforming the direct diagonalization of the Hill-Wheeler equations and performing equivalently to the symmetric orthogonalization results, though with much faster computation time.

We show results for the transition of $\hat{K}_x$ between the ground state and the first excited state. Once again, we see excellent agreement between trimmed sampling and the exact value, with similar accuracty to symmetric orhtogonalization. It drastically reduces the error from directly diagonalizing the Hill-Wheeler equation, and does so in much less computation time than is required for symmetric orthogonalization.

The results presented in this section were performed in a single execution of trimmed sampling, with each of the above observables calculated and weighted according to the posterior likelihoods for the matrix sets which produced them. In figure 3.14, we show the computation time for each value of the coupling. Trimmed sampling performs these calculations much faster than symmetric orthogonalization while producing similar results.
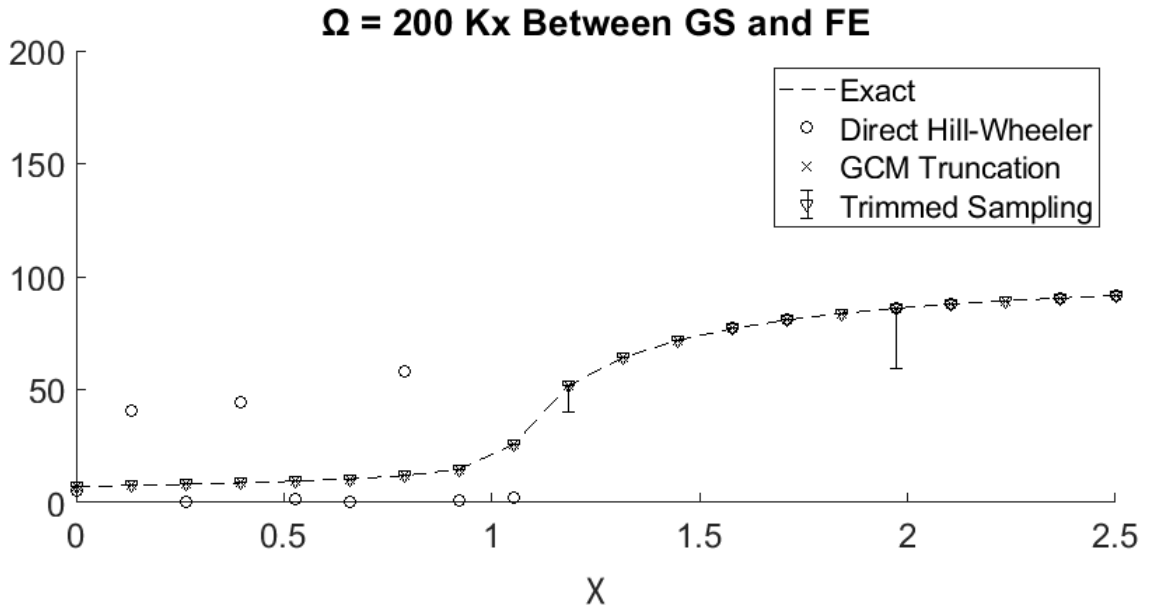
Figure 3.13: Matrix element of $\hat{K}_x$ between the ground state and first excited state. Symbols used are the same as in figure **??**. We see that both K-truncation and Trimmed Sampling produce similar results, while direct discretization of the Hill-Wheeler equations produces significant errors.
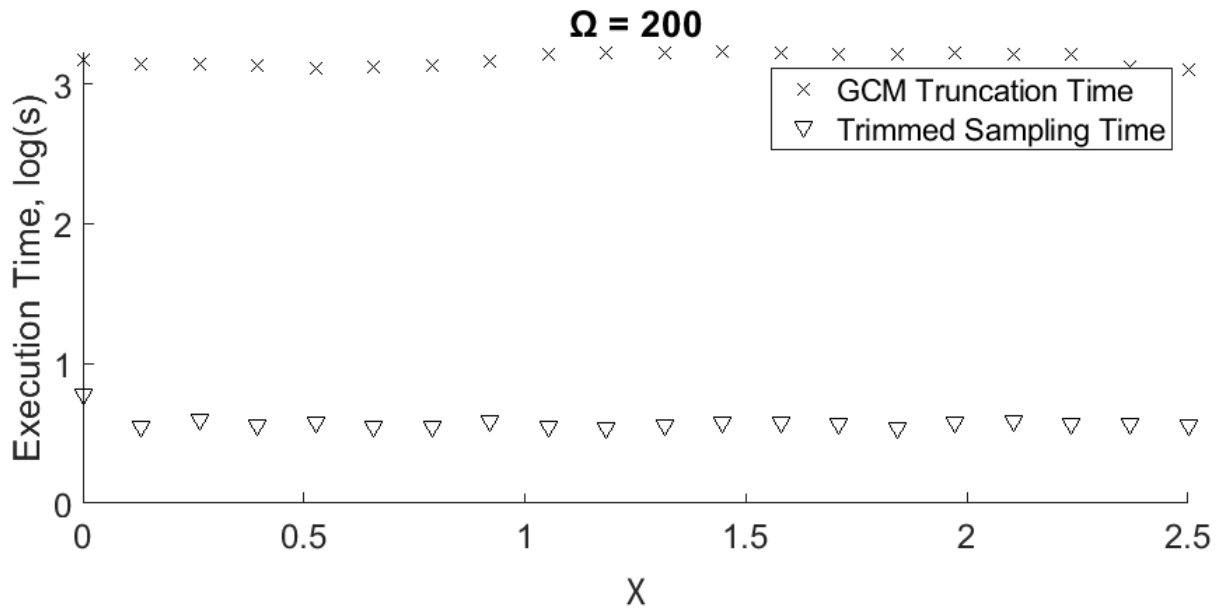
Figure 3.14: Calculation time for the Ground State Energy, First Excited State energy, and expectation values of the $\hat{K}_x$ operator for the k-truncation method and direct θ discretization using Trimmed Sampling. K-truncation is shown using 'x' markers, and trimmed sampling is shown using open triangle markers.

3.9.4 Signature Projection

For our second example of the Lipkin model, we will try a system much larger than in the first example. At this scale, symmetric orthogonalization becomes impractical. As a result, we will instead compare results to nugget regularization. We will also introduce error though truncation of matrix elements rather than an additional degenerate vector, and we will make use of different weight conditions.

3.9.4.1 Setup

In the previous example, one of our weight conditions, shown in equation 3.9.3, used the expectation value of $\hat{K}_x$ in the ground state as a posterior likelihood function. While there is good physical justification for this, in this example we will instead first project the system onto basis states which will ensure that $\langle \psi_0 | \hat{K}_x | \psi_0 \rangle = 0$ regardless of added noise. We accomplish this using signature projection.

In the Lipkin model, we define the signature operator as

$$R = e^{-i\pi \hat{K}_0}$$

The signature of an eigenstate is thus 1 if $\hat{K}_0$ is even, or -1 if $\hat{K}_0$ is odd. The Hamiltonian contains only terms which raise or lower the quasispin by 2, so these two sectors are entirely separate. We can use the projection operator,

$$\hat{P} = \frac{1}{2}(1 + r\hat{R})$$

to project a given state onto one of either positive signature (with $r = 1$) or negative signature (with $r = -1$). For a given state $|\tau\rangle$, we calculate its projection onto states of positive or negative signature as

$$\hat{P}_r |\tau\rangle = |\psi_r\rangle = \frac{1}{2}(|\tau\rangle + r|-\tau\rangle) \tag{3.46}$$

So for an arbitary operator, $\hat{O}$, we can calculate matrix values for states from sectors r and r' as

$$\langle \psi_r | \hat{O} | \psi_r' \rangle = \frac{1}{4} \left[\langle \tau | \hat{O} | \tau' \rangle + r' \langle \tau | \hat{O} | -\tau' \rangle + r \langle -\tau | \hat{O} | \tau' \rangle + rr' \langle -\tau | \hat{O} | -\tau' \rangle \right] \tag{3.47}$$

Using the result in equation 3.30, we can write the projected Norm operator as

$$N_{r,r'}(\tau, \tau') = \frac{1}{4} \left[\left[\cos(\frac{\theta - \theta'}{2}) \right]^\Omega + r' \left[\cos(\frac{\theta + \theta'}{2}) \right]^\Omega + r \left[\cos(\frac{\theta + \theta'}{2}) \right]^\Omega + rr' \left[\cos(\frac{\theta - \theta'}{2}) \right]^\Omega \right] \tag{3.48}$$

We can see that when $r = -r'$, the expression equals 0. We thus have two disconnected sectors, $r = r' = 1$ and $r = r' = -1$. Similarly, we can calculate the projected Hamiltonian matrix elements from equation 3.39, where we again see that we can divide our basis into two separate sectors, with $r = r' = 1$ and $r = r' = -1$.

We now apply equation 3.47 to calculate matrix elements of the $\hat{K}_x$ operator, using equation 3.43, to get

$$\langle\psi_r|\hat{K}_x|\psi'_r\rangle = \frac{1}{4}\frac{\Omega}{2}\left[\cos(\frac{\theta-\theta'}{2})^{\Omega-1}\sin(\frac{\theta+\theta'}{2}) - r\cos(\frac{\theta+\theta'}{2})^{\Omega-1}\sin(\frac{\theta-\theta'}{2}) + r'\cos(\frac{\theta+\theta'}{2})^{\Omega-1}\sin(\frac{\theta-\theta'}{2})\right.$$

(3.49)

Here, we now see that when $r = r'$, this equals 0, while it is nonzero only if $r = -r'$. The $\hat{K}_x$ operator connects the two sectors of positive and negative signature and is only nonzero when taken as the transition between two such states. As a result, by using these projected expressions, we can ensure that the expectation value of $\hat{K}_x$ in any eigenstate of the Hamiltonian is 0.

3.9.4.2 Updated weight criteria

The weight condition we used for the previous example, in equation 3.9.3, will no longer work as this condition will always be accomplished by default. However, this does come with a much small sampling space, as each sector has half the basis states as the system as a whole. We implement instead several new weight conditions. First, we use the same convergence criteria that we used for the energies but apply it to the transition of the $\hat{K}_x$ operator. In addition, we use a more robust method to determine which basis vectors to exclude for this convergence calculation. Rather than trimming vectors in a way which keeps the spacing between θ constant, we instead iteratively find the ground state for both the positive and negative signature sectors. We then remove the basis vector in each sector with the lowest overlap with the ground state. This process is more computationally expensive than just the prescribed method used in the last example, but because the ground state tends to be sharply peaked, removing equally spaced vectors could sometimes cause otherwise decent samples to be overly penalized if basis vectors with a lot of overlap with the ground state were removed.

Next, we use the signature of the ground state and first excited state as a filtering criteria. By construction, we expect one sector to have signature $r = 1$ and the other to have signature $r = -1$ for all eigenstates of the Hamiltonian. This provides a simple but effective filtering criteria

$$w_4 = e^{-a(\langle\psi_0|\hat{R}|\psi_0\rangle-1)^2 - a(\langle\psi_1|\hat{R}|\psi_1\rangle+1)^2}$$

(3.50)

where $|\psi_0\rangle$ and $|\psi_1\rangle$ are the ground states of the positive and negative sectors, respectively, and a is a parameter controlling the width of the distribution. This ensures that the matrix samples maintain signature as a good quantum number.
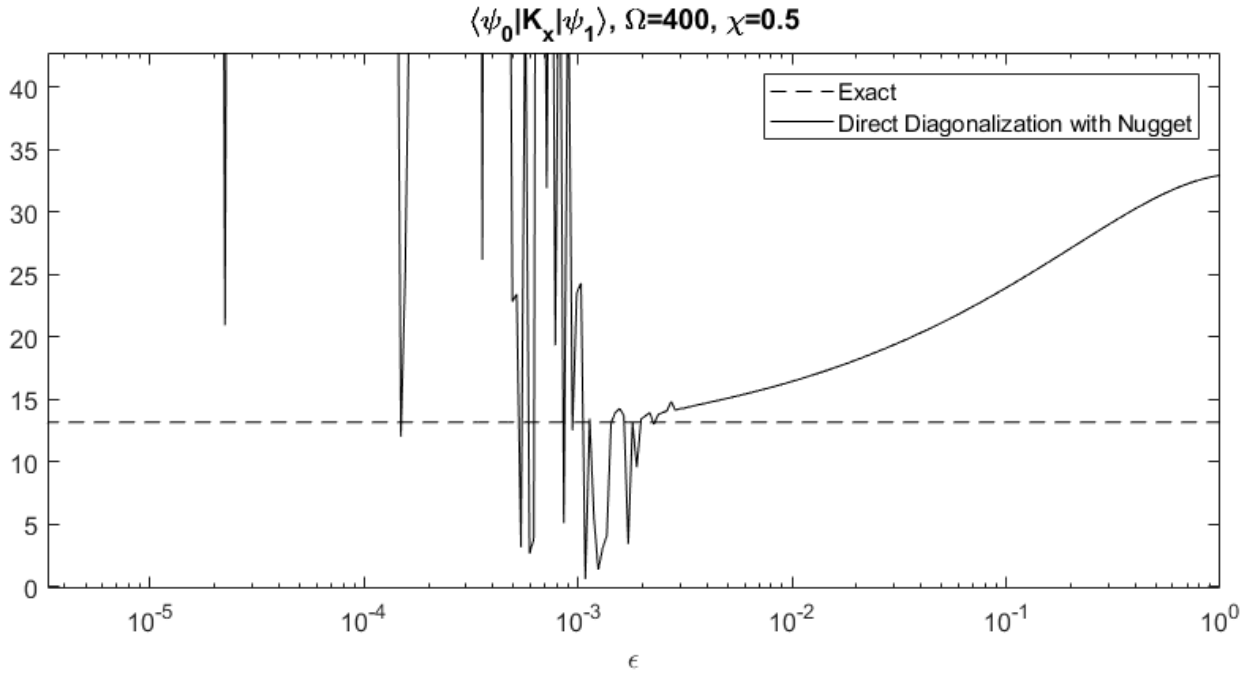
Figure 3.15: Possible values of $\hat{K}_x$ transition depending on nugget choice for $\Omega = 400$, $\chi = 0.5$, and 256 basis vectors. Each matrix element is rounded to the third decimal to introduce error and clearly show the instability. The exact solution is shown with a horizontal dotted line, while the value of $\langle\psi_0|\hat{K}_x|\psi_1\rangle$ is shown with a solid line. We see that while large values of the nugget produce increasingly smooth behavior, it does so at the cost of drastically impairing the accuracy of the calculation. Smaller values of the nugget are insufficient to correct the instability.

3.9.4.3   Nugget Regularization

As we will be testing systems at a scale which makes symmetric orthogonalization impractical, we will instead be comparing results to nugget regularized results. We add a nugget, a small multiple of the identity, ϵI, to the diagonal of the norm matrices. This helps to reduce the instability due to the poor conditioning of the norm matrices. However, it can be difficult to determine an appropriate value for ϵ, and it is not easy to determine the systematic effects this value has on an observable in question. As we will see in this example, in some cases no value of the nugget is sufficient to correct the system's instability.

Figure 3.15 shows a wide range of possible values for the nugget parameter, ϵ, for a single value of the coupling, $\chi = 0.5$, and $\Omega = 400$. In this case, we see that at small values of ϵ, the value calculated is both far from the exact value and highly dependent on the value of ϵ. As ϵ increases, the system eventually becomes more stable with respect to varying ϵ, but by this point the value of the nugget is already large enough to significantly altar the value we are attempting to measure. As a result, there is no workable value of ϵ which we can rely on to correct the instability. As we will see, however, trimmed sampling is able to successfully correct this error.

3.9.4.4   Results

Using the same parameters we described in the previous section, we now apply trimmed sampling to this case of the Lipkin model. For this example, we use 256 basis vectors and introduce error by truncating each matrix element at the third decimal. As we showed in figure 3.15, there is not a good value for the nugget which provides both good error correcting and is stable with respect to the choice of the nugget parameter, ϵ. As a result, we chose a value of 10^{-3}, which is slightly in the unstable region but still provides reasonably accurate results at higher values of the coupling.

In figure 3.16, we show the results for the ground state, which is the ground state in the sector with positive signature. We see that nugget regularization fails to fully correct the error, especially at low values of the coupling. Trimmed sampling provides excellent correction for this case.

We see similar results for the first excited state, which is the ground state of the negative signature sector. Figure 3.17 shows these results. In this case, we see that the nugget regularization is again unable to fully correct the error, especially at low values of the coupling. Trimmed sampling does not have this problem and provides excellent error mitigation at all values of the coupling.

In figure 3.18, we show the value of the $\hat{K}_x$ transition between the ground state and first excited state, which are the two lowest energy states in their signature sectors. We see that nugget regularization is completely unable to correct the error for low values of
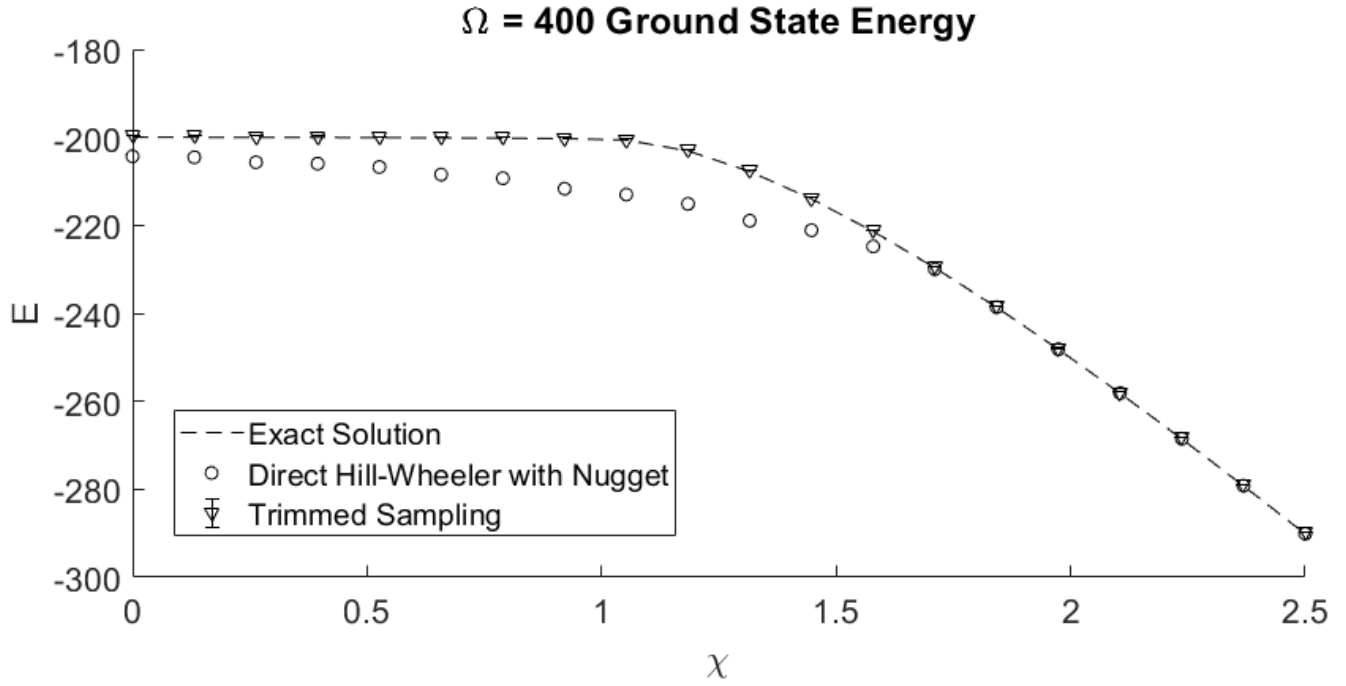
Figure 3.16: Ground state energy in the Lipkin model with $\Omega = 400$, using 256 basis vectors to discretize the Hill-Wheeler equations. This is the ground state from the positive signature sector. The dashed line shows the exact values. Results in open circles show values after the norm matrix is regularized using a nugget ϵI with $\epsilon = 10^{-}3$. This value was chosen as it provides a good level of correction and results are not too sensitive to the value of ϵ near this value. Triangle markers show the ground state energy after applying trimmed sampling. We see much better performance from trimmed sampling than we do from nugget regularization.
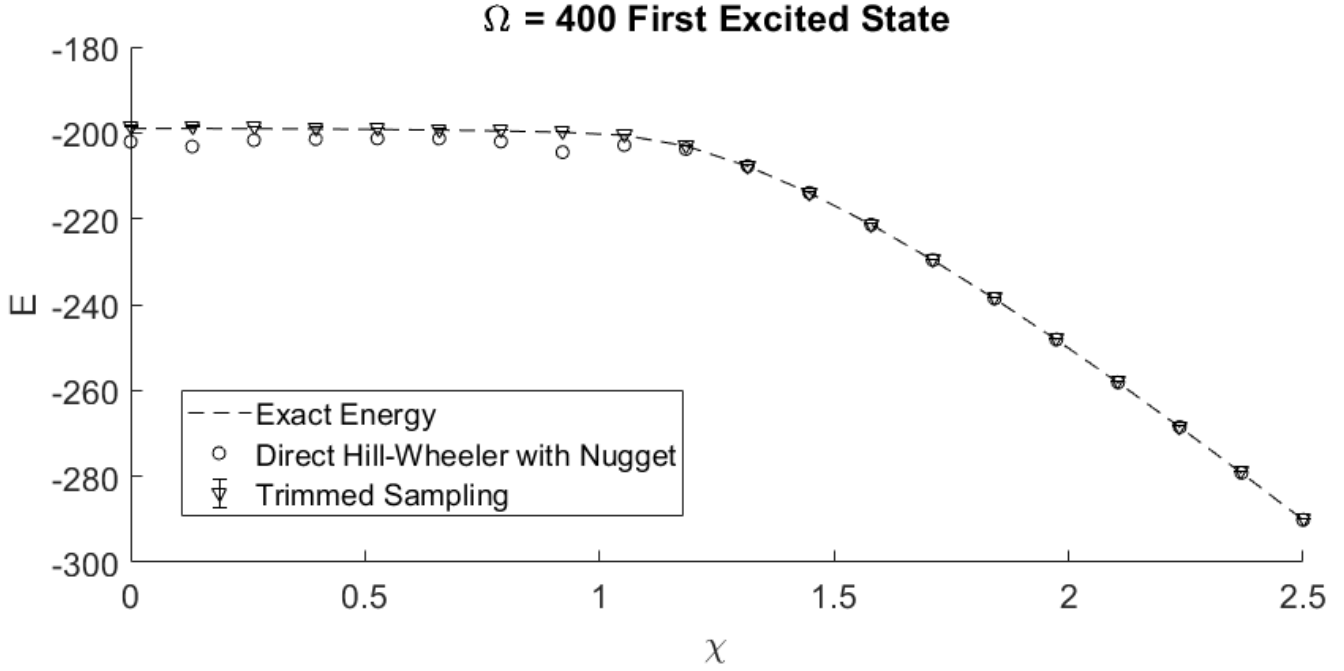
Figure 3.17: First excited state energy of the Lipkin model with $\Omega = 400$, using 256 basis vectors. This is the ground state from the negative signature sector. Symbols are the same as figure 3.16. As with the ground state, we see that trimmed sampling performs much better than the nugget regularization.
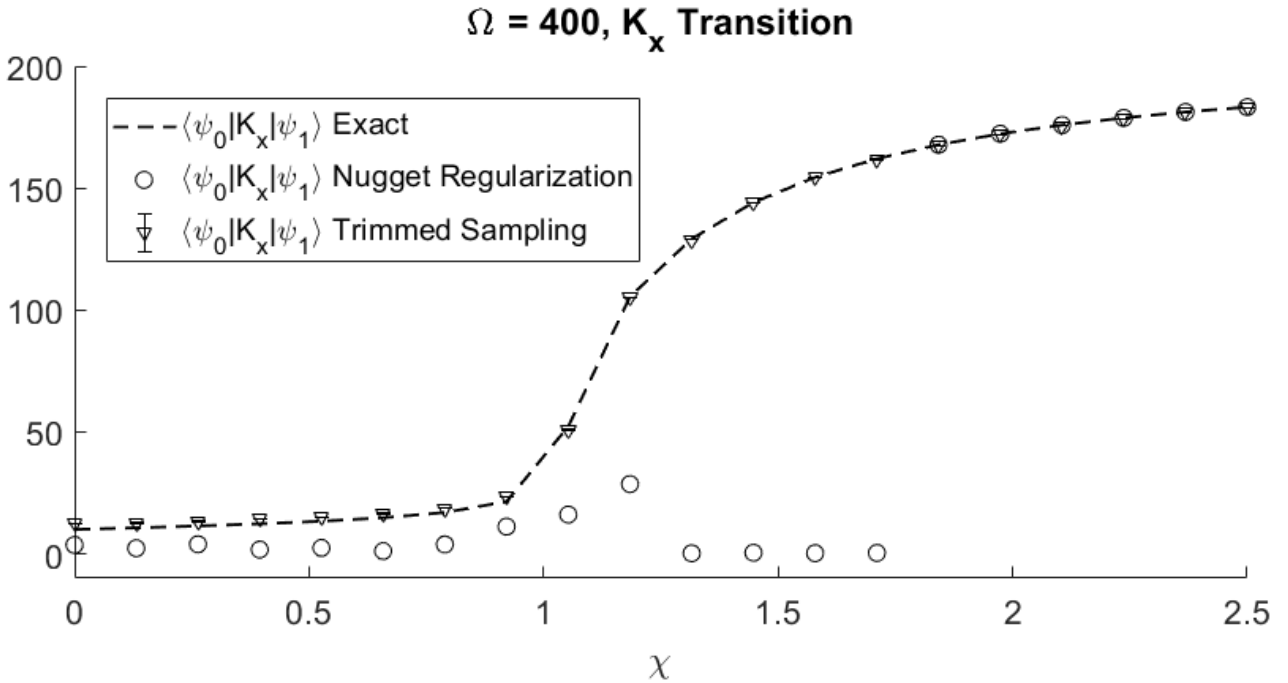
Figure 3.18: Transition strength of $\hat{K}_x$ between the ground state, $|\psi_0\rangle$, and first excited state, $|\psi_1\rangle$, which are the lowest energy states in their respective signature sectors. The exact value is shown as a dotted line. Results using nugget regularization are shown with open circles. Trimmed sampling results are shown with triangle markers.
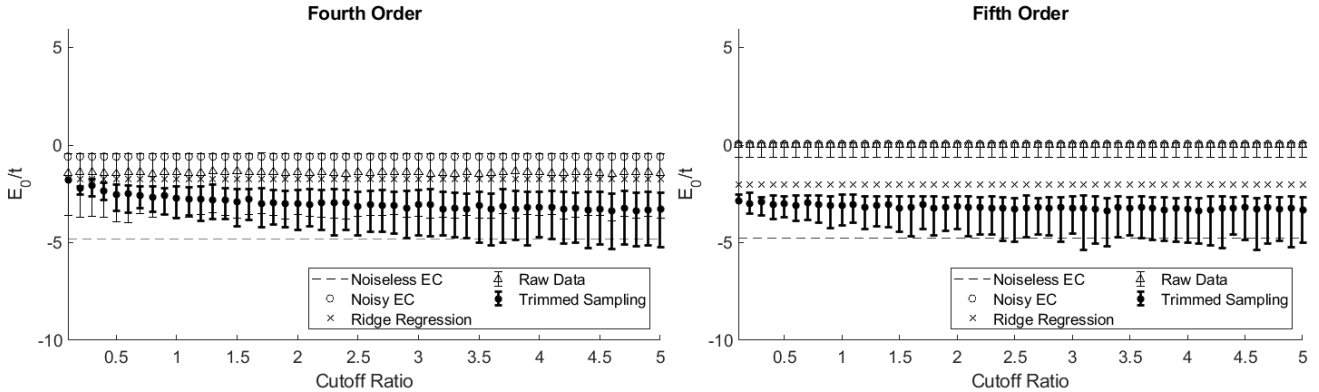
Figure 3.19: Comparison of various choices for the cutoff ration, C, for the Bose-Hubbard model example. Results for fourth order calculations are presented in the left panel, and results for fifth order calculations are presented in the right panel.

the coupling, while trimmed sampling performs well for all values of the coupling and is able to correct the error to values much closer to the exact values.

At large system size, symmetric orthogonalization becomes impractical due to the computational scaling. Trimmed sampling and nugget regularization both remain applicable at large system scales, but nugget regularization is unable to fully correct errors in this case. We see that the nugget choice can have drastic effects on the results, and in some cases no value of the nugget provides good error correction. Trimmed sampling is able to correct the error in this cases much better than nugget regularization, and remains fast and practical well beyond the point where symmetric orthogonalization becomes impractical.

3.10   Non-Sensitivity to User Choice

To ensure that the trimmed sampling results are not biased by user choice, we present here several sensitivity studies. The choices for various parameter and functional forms in the previous examples were selected due to simplicity, but they might appear arbitrary. We will show here that the final results are not sensitive to these choices.

3.10.1   Cutoff Ratio

We will begin by looking at the sensitivity of our commonly used filtering criteria, fig. 3.7. We examine the case in the Bose-Hubbard example where the coupling strength $U/t = -4$. We chose this coupling strength as it lies beyond the avoided level crossing and we observe large differences between the methods presented. We vary the convergence ratio parameter, C, over the range from 0.2 to 5. These results are plotted in fig. 3.19. We see that, provided the convergence ratio is not overly restrictive, the value of the

Figure 3.20: Calculation of energies for the Bose-Hubbard model using a Gaussian form for the likelihood function. Results for fourth order calculations are presented in the left panel, and results for fifth order calculations are presented in the right panel.

cutoff ratio has only limited impact on the results. In cases where C is very small, we are systematically biasing the posterior probability so much that the reported error bars are significantly smaller than deviation from the exact result. However, when the value of C is not very restrictive, we see that the result does not vary, and the error bars correctly capture the actual deviation from the exact result.

3.10.2 Functional Form

The functional form in fig. 3.7 was selected due to its simplicity, but several other potential functional forms may appear equally valid. We show here several example functional forms and see that while very poorly suited functions may produce noticeably worse results, all results still show substantial improvement over ridge regression and the noisy case. For each of these examples, parameters match those presented previously for the Bose-Hubbard model solved using eigenvector continuation, with the only change being the functional form of the cutoff for the convergence ratio.

In figure 3.20, we show results using a Gaussian cutoff function of the form

$$f_C(H, N) = e^{-\frac{C_{max}^2}{C^2}}$$

The exact normalization of these functions is unimportant as these normalizations cancel out when we take the weighted median. We see that the 14th and 86th percentile marks shift somewhat relative to the exponential form presented previously, but the results are qualitatively the same. The Gaussian form goes to 0 much faster than the exponential form for large C_{max}, so a narrower distribution of matrix pairs is accepted, and so the reported error bars are somewhat smaller than the deviation from exact results.

In figure 3.21, we show the results from using a linear-times-exponential function with the form

$$f_C(H, N) = C_{max} e^{-\frac{C_{max}}{C}}$$

Figure 3.21: Calculation of energies for the Bose-Hubbard model using a linear-times-exponential form for the likelihood function. Results for fourth order calculations are presented in the left panel, and results for fifth order calculations are presented in the right panel.
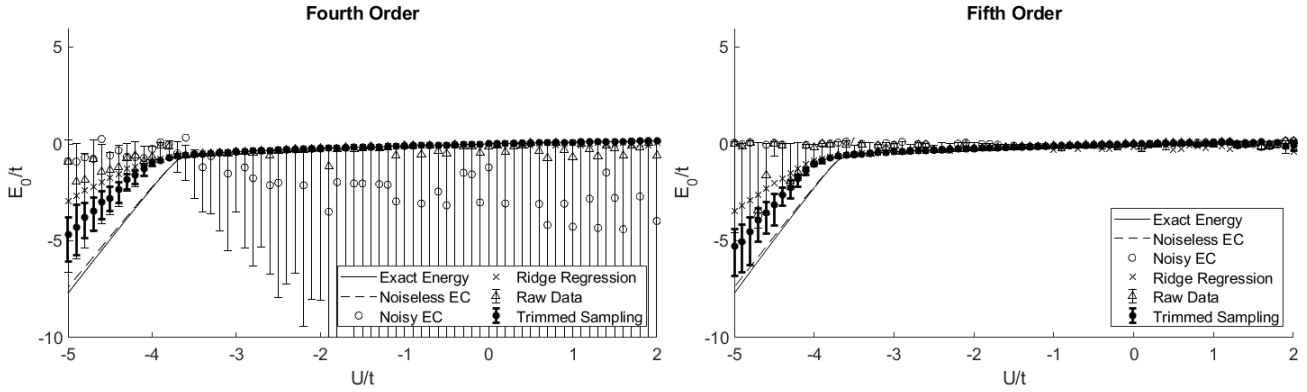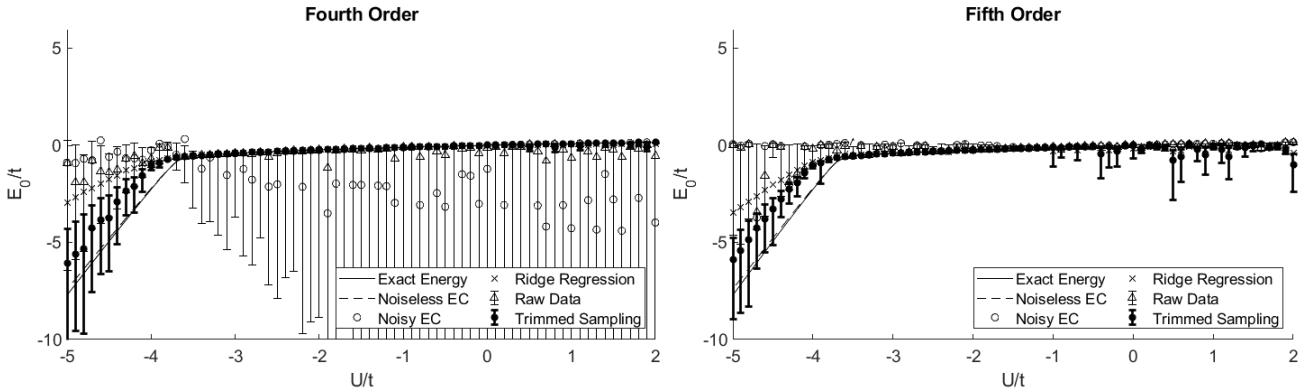

Figure 3.22: Calculation of energies for the Bose-Hubbard model using a rational function form for the likelihood function. Results for fourth order calculations are presented in the left panel, and results for fifth order calculations are presented in the right panel.

Again in this case, we see that the 14th and 86th percentile marks have shifted, in this case they have widened, but the qualitative results are still very similar to the exponential case. This functional form has a peak at C rather than 0, as the other functions presented here do. Its behavior is somewhat worse than in previous examples, but still shows substantial improvement over the noisy case and ridge regression. This functional form makes little sense in this context, but is presented to show that even a poorly suitd functional form still shows substantial improvement.

In figure 3.22, we show the results of the Bose-Hubbard calculation again, this time using a functional form for the cutoff of a reciprocal linear function,

$$f_C(H, N) = \frac{1}{1 + \frac{C_{max}}{C}}$$

The results from this test are comparable to the others presented, but as this function is not exponentially damped at large values of C_{max} the error bars indicate a much wider
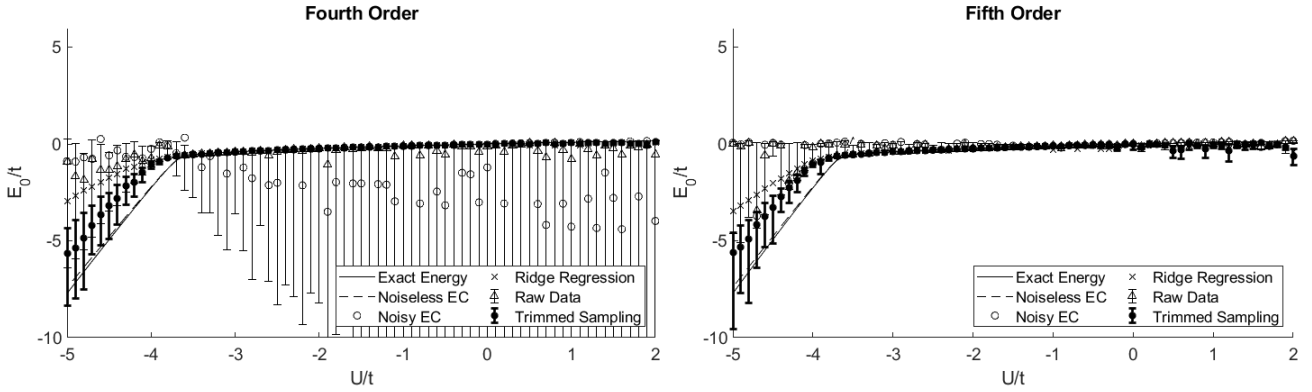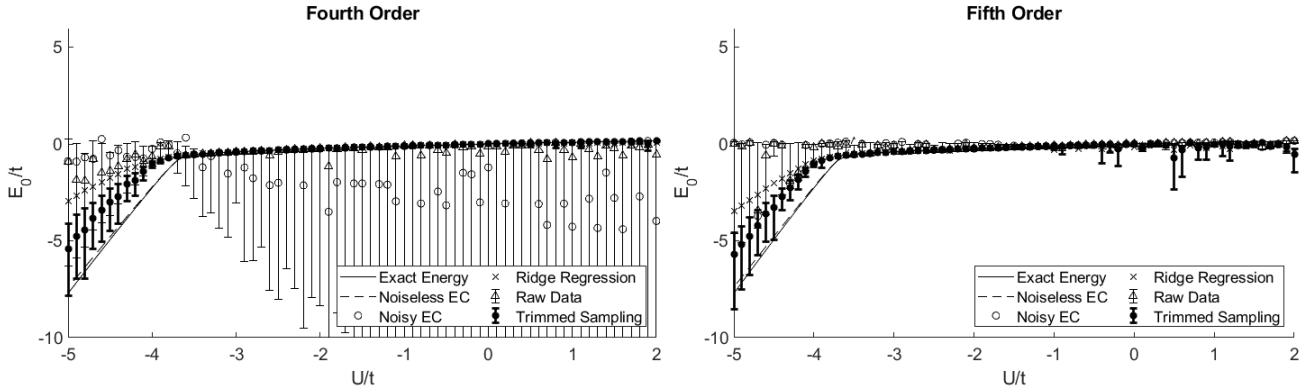
Figure 3.23: Calculation of energies for the Bose-Hubbard model using a rectified sinc function form for the likelihood function. Results for fourth order calculations are presented in the left panel, and results for fifth order calculations are presented in the right panel.

posterior distribution. It still shows improvement over the noisy EC and reidge regression. Finally, in figure 3.23, we use a rectified sinc function likelihood facter of the form

$$f_C(H, N) = \max \left(\frac{\sin \frac{C_{max}}{C}}{C_{max}}, 0 \right)$$

This function form has no physical justification in this case, and is presented to show that even with very poor choice of functional form we still see significant improvement over the noisy EC and ridge regression.

While the function form of the likelihood function does have an effect on the results, in particular on the edges of the posterior distribution, we see that the effect of user choice on the likelihood function is minimal, and still allows trimmed sampling to perform well despite particularly poor user choices. In some examples, we use varying functional forms, but as the results are not impacted significantly these functional forms are typically chosen for ease of sampling. For the Eigenvector Continuation example on the Bose-Hubbard model, we chose the exponential form because it is simple, not overly restrictive, exponentially damps large values of C_{max}, and is easy to sample.

3.11 Conclusion

Trimmed sampling provides a simple and scalable approach to error mitigation of noisy generalized eigenvalue problems. By weighting sample systems by physics informed likelihood functions, we are able to determine values of a given observable are possible and likely within our given physics constraints. Understanding this distribution allows us to significantly reduce the effects of errors on sensitive systems and can have significant applications in any noise sensitive inverse problem.

In the first model, we were able to correct an extremely sensitive application of eigenvector continuation to the Bose-Hubbard model. While even tiny error relative to the

matrix elements was sufficient to render the results for the ground state energy unusable, we were able to use extremely simple filtering criteria with trimmed sampling to correct the majority of the error and get a sense of the range of possible physical values.

In the second example, we introduced noise to the generalized eigenvalue problem produced by solving the Heisenberg chain model using Euclidean time projection. Despite this noise again causing the results to become unusable, we were able to recover the correct behavior and produce accurate results. We were also able to extend the method to calculations of a chosen spin observable, showing that trimmed sampling is reconstructing not just the eigenvalues of the generalized eigenvalue problem but also its eigenvectors.

Finally, in the third example we showed that trimmed sampling can be applied at a much greater scale than symmetric orthogonalization and performs better than ridge regression by applying trimmed sampling to solutions of the Lipkin model produced using generator coordinate method. By implementing several improvements to the original algorithm, we were able to increase its convergence time substantially and produce strong results even on extremely large and sensitive generalized eigenvalue problems.

Trimmed sampling is an algorithm which allows us to sample the distribution of physically realizable states that can emerge from a generalized eigenvalue problem. By sampling and understanding this distribution, we are able to substantially reduce the effects of noise which may otherwise render the results unusable.

4

Projected Cooling

4.1 Overview

The quantum many-body problem is a challenge which arises in nearly all branches of quantum physics with far reaching applications, such as *ab initio* methods for nuclear structure and reactions (40, 41), strongly coupled electrons, degenerate atomic gassis, and interacting relativistic quantum fields. Many of these problems are difficult to treat on a classical computer architecture, but quantum computing provides a new computing paradigm with the potential to overcome the significant difficulties present in classical calculations of the quantum many-body problem.

Quantum algorithms allow arbitrary superpositions of products of qubits, allowing them to store exponentially more information than classical bits without requiring stochastic sampling, alleviating some of the difficulty when scaling classical simulations to large quantum systems. However, this increased efficiency currently comes at the cost of short decoherence times and significant readout errors. To mitigate these errors, we must develop quantum algorithms which are robust against noise without.

Many methods currently exist to calculate the ground state of a given Hamiltonian on a quantum computer, such as quantum phase estimtimation (42, 43), quantum variational methods(27, 44), quantum adiabatic eveolution(45, 46), spectral comb techniques(47), resonance transition enhancement(48), coupled heat bath approaches(49, 50), and dissipative open-system methods(51, 52). Each of these methods have difficulty reaching reliable accuracy for Hamiltonians of interest. We therefore introduce the projected cooling algorithm to address the need for an efficient and accurate method to calculate the ground state of a Hamiltonian of interest.

Projected cooling works as a quantum analog of evaporative cooling. Rather than beginning with hot gas molecules, we begin with some initial state $|\psi_I\rangle$ which has cupport over a compact region, called ρ. We then drive the excited quantum states to disperse away from ρ, leaving only the lowest-lying states behind in the compact region. We can then measure the remaining wave function left behind in ρ. In this way, projected

cooling can construct the localized ground state of any Hamiltonian with a translationally-invariant kinetic energy provided all interactions go to zero at large distances.

We will consider three models which we call 1A, 1B, and 2.

4.2 Model 1

For models 1A and 1B, we begin with a Hamiltonian defined on a one dimensional chain of 2L+1 qubits labeled $n = -L, \cdots, L$. We consider the vacuum state to be the state where all qubits are in the $|0\rangle$ state and define particle excitations as qubits in the $|1\rangle$ state. For example, if qubit n is in the $|1\rangle$ state, then we can say that there is a particle at site n. This is analogous to a spatial lattice formalism which is commonly used in classical computing methods, such as lattice effective field theory calculations of nuclear and cold atomic systems(53, 54). In this language of second quantization, the number of particles is equal to the number of $|1\rangle$ qubits. However, as all systems we will consider have particle-number conserving Hamiltonians, it will be more convenient to use the language of first quantization where particle number is fixed and to use spatial wavefunctions of such states.

4.2.1 Model 1A

Let us consider the one particle subspace, where $|[n]\rangle$ is the tensor product state where qubit n is $|1\rangle$ and all other qubits are in the $|0\rangle$ state. In this subspace, we define our Hamiltonian as $H = K + V$ with $\langle [n']|H|[n]\rangle = K_{n',n} + V_n \delta_{n',n}$, where $\delta_{i,j}$ is the Kronecker delta function. The kinetic energy term $K_{n',n} = \delta_{n',n} - \frac{1}{2}\delta_{n',n+1} - \frac{1}{2}\delta_{n',n-1}$ is derived from the finite difference and V_n is the single-particle energy on site n. For model 1A, we will consider a 'spiked' attractive potential, with $V_n = -\delta_{0,n}$. We define a compact region ρ to correspond to the qubits $n = -R, \cdots, R$ with $R \ll L$. The compact region ρ must have a volume that is small compared to the volume of the total system. We now define the operator P as the projection operator which projects a given state onto the subspace where all particle excitations are contained entirely within ρ. P can be constructed explicitly as a product of $|0\rangle\langle 0|$ over all qubits outside of ρ, so that $P|[n]\rangle = 0$ for $|n| > R$ and $P|[n]\rangle = |[n]\rangle$ for $|n| \leq R$. This will allow us to eliminate the higher energy states which escape from the compact region.

Let $|\psi_0\rangle$ be the ground state of H. For model 1A, the ground state, $|\psi_0\rangle$, is localized in position space, bound, and is the only bound state of H. All other states are continuum states and so extend to infinity in the limit $L \to \infty$. We use dimensionless quantities and set $\hbar = 1$, so our time evolution operator is defined as $U(t) = e^{-iHt}$. Projected cooling works because in the limit $L \to \infty$, the projected time evolution operator $PU(t)P$ has a stable fixed point proportional to $P|\psi_0\rangle$. As the time operator $U(t)$ acts on an arbitrary
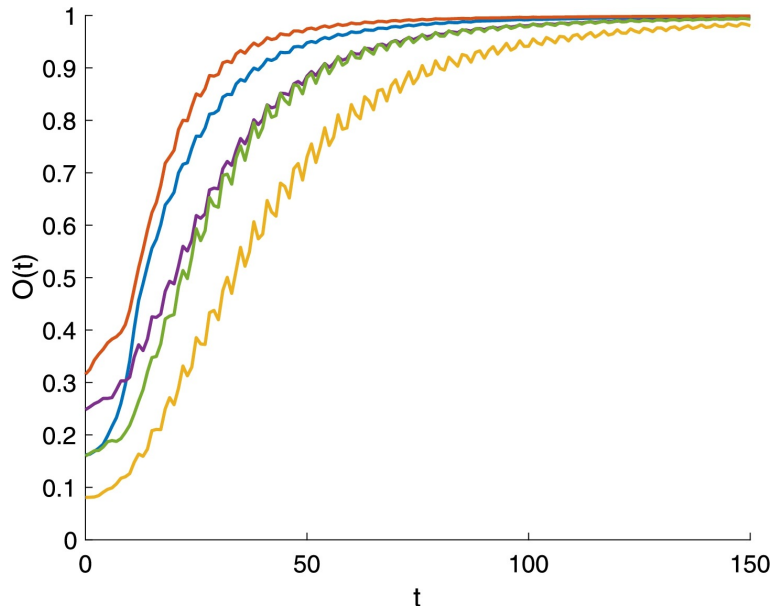
Figure 4.1: Stable fixed point for Model 1A. The normalized overlap $O(t)$ is shown between the exact wave function and the evolved wave function over the interior region ρ versus time t for model 1A. These results show five random initial states, and we see that all approach $O(t) = 1$ for large t.

state $P|\psi_I\rangle$, excited continuum states leave the compact region ρ and are projected out. In the limit of large t, therefore, the only part of the wavefunction which remains when projected by P is from the bound state $|\psi_0\rangle$. This assumes that the initial state is not orthogonal to the bound state wavefunction, in which case the signal would go to 0 at large t. This behavior is exhibited by any Hamiltonian with translationally-invariant kinetic energy, interactions which vanish at large distance, and exactly one localized bound state. If $P|\psi_I\rangle$ has well-defined quantum numbers associated with an exact symmetry of H, then the fixed point property applies to cases where exactly one localized bound state exists within the symmetry sector of interest.

To evaluate the success of the method, we define a normalized overlap of two states $|x\rangle$ and $|y\rangle$ to be $|\langle x|y\rangle / \sqrt{\langle x|x\rangle\langle y|y\rangle}$. Let $O(t)$ be the normalized overlap between $P|\psi_0\rangle$ and $PU(t)P|\psi_I\rangle$. In figure 4.1, we show this normalized overlap $O(t)$ for five random initial states, $|\psi_I\rangle$, versus time t for model 1A. These states are chosen randomly to demonstrate that the fixed point behavior is universal and not overly dependent on initial conditions. For this example, we take $R = 5$ with a large enough box to prevent reflection waves returning from the boundary. In all cases, the normalized overlap approaches $O(t) = 1$, signifying that the evolved wave functions approach $|\psi_0\rangle$ at large t and that $P|\psi_0\rangle$ is a fixed point.

We apply the projection operator P only at the end of the time evolution, and the probability of measuring a non-zero signal is determined by $|\langle\psi_0|P|\psi_I\rangle|^2$, the squared overalap of the initial and ground states. This signal efficiency is independed of the

projection time and relies on the quality of the initial wave function. For this reason, smooth initial wave functions, such as a Gaussian wave packet, in the compact region are often reasonable choices.

Projected cooling is also able to calculate a given observable of interest, O. Let us say that O_p is the same observable restricted to the region of interest, ρ. The projection operator, P, amounts to simultaneously measuring the Pauli-Z operator on each qubit outside the inner region, ρ, with each qubit collapsing to either $|1\rangle$ or $|0\rangle$. We can then record the measurement of O_p if and only if we find that each exterior qubit has collapsed to the $|0\rangle$ state. In systems with many particle excitations, we can instead record the measurement of O_p when the number of exterior $|1\rangle$ qubits is lower than some small number, δ, times total particle number.

4.2.2   1B

In cases where we wish to apply projected cooling to a Hamiltonian with more than one bound state, we can make use of adiabatic evolution. If we are interested in the ground state of a Hamiltonian H, we can define a time dependent Hamiltonian $H(t)$ such that its initial stage is a Hamiltonian with only one bound state, but which has good overlap with the ground state of the final Hamiltonian H. We can achieve this by scaling the kinetic energy operator by a factor greater than 1, until all bound states are pushed into the continuum. This augmented kinetic energy operator also accelerates the time evolution at early times, requiring fewer quantum gate operations per qubit. Once projected cooling has produced a state $|\psi(t)\rangle$ which closely tracks the ground state of $H(t)$, we can use adiabatic evolution to gradually reach the ground state of our desired Hamiltonian, H.

We will now consider model 1B, defined in the same way as model 1A but we now take $V_n = -1.6\delta_{0,n} - 1.5(\delta_{2,n} + \delta_{3,n}) - 1.4\delta_{-2,n}$. This change produces a Hamiltonian with 4 bound states. We will first compute the results using the full time evolution operator, $U(t, t-\Delta t) = e^{-iH(t)\Delta t}$. We can also use the Trotter approximation to break the Hamiltonian into pieces, $H = A + B + D + V$ where $A_{n',n}$ is the off-diagonal component of $K_{n',n}$ when $\min(n', n)$ is even, $B_{n',n}$ is the off-diagonal component of $K_{n',n}$ when $\min n', n$ is odd, and $D_{n',n}$ is the diagonal part of $K_{n',n}$. The Trotterized time evolution operator is thus $e^{-iA(t)\Delta t}e^{-iB(t)\Delta t}e^{-iD(t)\Delta t}e^{-iV(t)\Delta t}$. To show the effect of stochastic noise on the time evolution, we multiple each component of the evolved wavefunction by a factor $1 + z$ after each time step, Δt, where z is a complex Gaussian random variable centered at zero with root mean square values of $\epsilon/\sqrt{2}$ for real and imaginary parts.

For the adiabatic evolution, we start with an initial Hamiltonian $H_I = V$. We take the ground state of H_I, a point-like wave function $|\psi_I^1\rangle = |[0]\rangle$, as our initial state. For the projected cooling calculation, the initial state has more freedom and can be any state contained within the region ρ. For projected cooling, we use the same point-like wave
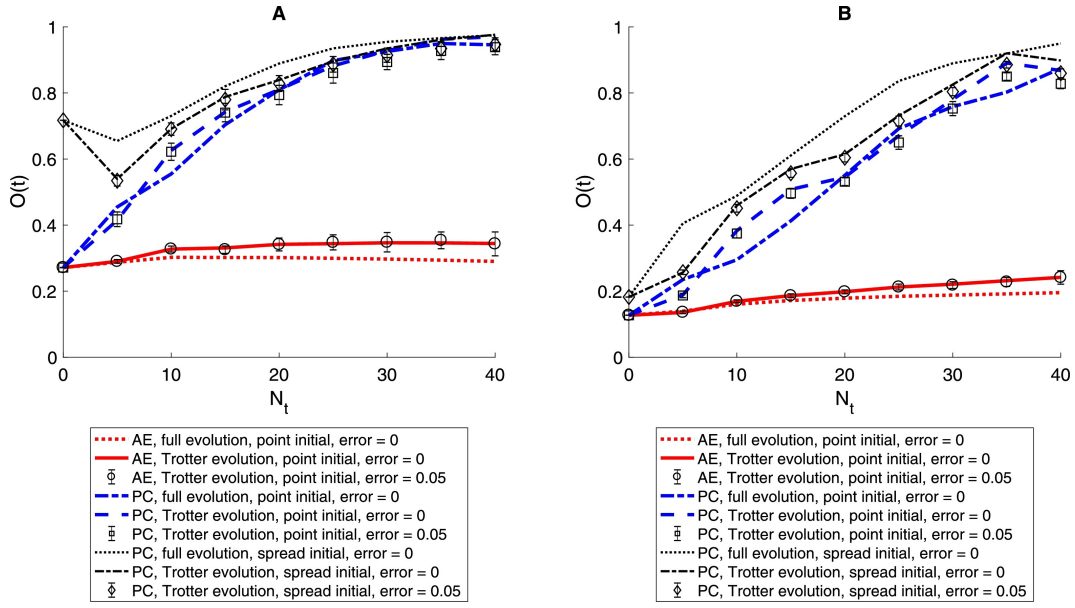
Figure 4.2: We show the normalized overlap $O(t)$ between the evolved wave function and the exact wave function over the interior region ρ versus the number of time steps N_t. Results from model 1B are on the left, and results from model 2 are on the right. AE is adiabatic evolution and PC is projected cooling. Full evolution denotes time evolution using the full time-dependent Hamiltonian at each step, while Trotter evolution denotes using the Trotter approximation. Point initial indicates that the initial state is the point-like initial state, $|\psi_I^1\rangle$, while spread initial indicates that the initial state is the spread out state $|\psi_I^2\rangle$. The quoted error is the value of the parameter ϵ in the noise added.

function in addition to the smeared wave function $|\psi_I^2 = 0.75|[0]\rangle + 0.43(|[1]\rangle + |[-1]\rangle) + 0.26(|[2]\rangle + |[-2]\rangle)$. We show the normalized overlap $O(t)$ between the evolved wave function and the exact wave function in the interior region ρ versus the number of time steps, N_t, for model 1B in the first panel of figure 4.2. For this test we take $R = 5$, $L = 25$, and the time step $\Delta t = 0.3$. This corresponds to an 11 dimensional inner region, ρ, in the one-particle sector.

We see that Adiabatic evolution has significant difficulties with finding the ground state and achieves no greater than 0.35 overlap with the target ground state. Projected Cooling performs much better, achieving an overlap of at least 0.94 in 40 or fewer time steps for all cases, even when errors due to the Trotter approximation and stochastic noise of size $\epsilon = 0.5$.

4.3 Model 2

For Model 2, we consider a Hamiltonian defined on two linked one-dimensional chains of 2L+1 qubits each, $n_1 = -L, \cdots, L$ and $n_2 = -L, \cdots, L$. The vacuum state is defined in the same was as with model 1, as the tensor product state where all qubits are $|0\rangle$. We will consider the two-particle sector and define $|[n_1, n_2]\rangle$ as the tensor product state where qubin n_1 on the first chain is $|1\rangle$ and qubit n_2 on the second chain is also $|1\rangle$, with all other qubits as $|0\rangle$. We define the interior region ρ as the sites where $\max(|n_1|, |n_2|) \le R$ with the values $R = 5$, $L = 25$, and $\Delta t = 0.3$. We now have an interior region with 121 dimensions in the two-particle sector.

The Hamiltonian has the form $H = K + V + W$ with

$$\langle [n_1', n_2']|H|[n_1, n_2]\rangle = K_{n_1', n_1}\delta_{n_2', n_2} + K_{n_2', n_2}\delta_{n_1', n_1} + (V_{n_1} + V_{n_2} + W_{n_1, n_2})\delta_{n_1', n_1}\delta_{n_2', n_2}$$

where the kinetic energy term $K_{n', n}$ is the same as in Model 1. For the interactions, we take $V_n = -1.0\delta_{0,n} + 0.2\delta_{1,n} - 0.9(\delta_{n,n} + \delta_{3,n}) - 0.3\delta_{-1,n}$ for the single particle energy and $W_{n_1, n_2} = -0.2\delta_{n_1, n_2}$ for the two-particle interaction. This model has four localized bound states which remain within the region ρ. For adiabatic evolution, we will use the point-like initial state $|\psi_I^1\rangle = |[0, 0]\rangle$. For projected cooling, we consider both $|\psi_I^2\rangle$ and the smeared initial state

$$|\psi_I^2\rangle = 0.81|[0, 0]\rangle + 0.30\left(|[1, 0]\rangle + |[-1, 0]\rangle + |[0, 1]\rangle + |[0, -1]\rangle\right)$$

For the Trotter approximation, we split the Hamiltonian into the pieces

$$H = A^{[1]} + B^{[1]} + A^{[2]} + B^{[2]} + D + V + W$$

where $A^{[i]}_{n_i', n_i}$ is the off diagonal part of the kinetic energy for particle i when $\min(n_i', n_i)$ is even, $B^{[i]}_{n_i', n_i}$ is the off-diagonal part of the kinetic energy of particle i when $\min(n_i', n_i)$
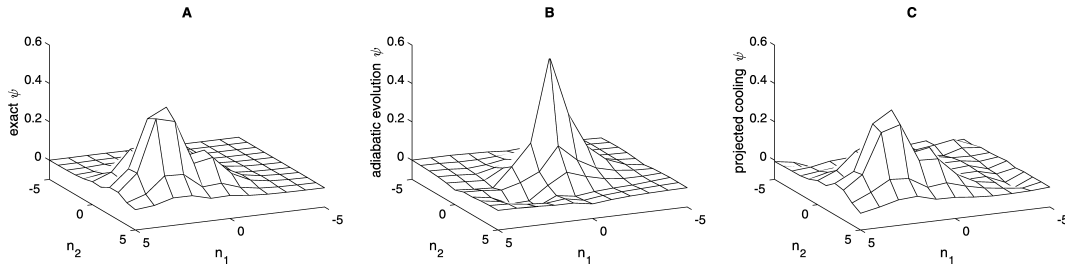
Figure 4.3: The exact ground state wave function in the interior region of Model 2 is shown in the left most panel, A. Panel B shows the wave function obtained via 40 time steps of adiabatic evolution. The right panel, C, shows the wave function obtained by projected cooling after 40 time steps. For each of these, we use $R = 5$, $L = 25$, and the time step $\Delta t = 0.3$.

is odd, and D is the diagonal part of the kinetic energy. We will call the static, time-independent operators as $\bar{H}, \bar{K}, \bar{V}$ etc.

For the adiabatic evolution from initial time $t = 0$ to final time $t = t_f$, we use the time dependent Hamiltonian

$$H(t) = \frac{t}{t_f}\bar{K} + \bar{V} + \bar{W}$$

For the projected cooling, we use the time dependent Hamiltonian

$$H(t) = (10\bar{K} - \bar{H})\exp(-t/3.6) + \bar{H}$$

and the Trotterized time evolution operator is

$$U(t, t - \Delta t) = e^{-iA^{[1]}(t)\Delta t}e^{-iB^{[1]}(t)\Delta t}e^{-iA^{[2]}(t)\Delta t}e^{-iB^{[2]}(t)\Delta t}e^{-iD(t)\Delta t}e^{-iV(t)\Delta t}e^{-iW(t)\Delta t}$$

The right panel of figure 4.2 shows results of the normalized overlap $O(t)$ between the evolved wave function and the exact wave function in the interior region ρ versus the number of time steps, N_t, for Model 2. As with model 1B, we see that standard adiabatic evolution has difficulties finding the ground state, and does not achieve an overlap greater than 0.24. Projected cooling provides a significant improvement, achieving an overlap of at least 0.85 in 40 time steps or less for all cases, even with errors from the Trotter approximation and stochastic noise of size $\epsilon = 0.5$.

In Figure 4.3, we show a direct comparison of the wafe functions obtained by projected cooling and adiabatic evolution to the exact ground state in the interior region of Model 2. For these we use the spread initial state $|\psi_I^2\rangle$ for projected cooling, full time evolution, and error $\epsilon = 0$. The wave function obtained using projected cooling is significantly closer to the exact wave function than the wave function obtained using adiabatic evolution.

4.4   Conclusions

The results presented from these models show significant improvement over other methods, such as adiabatic evolution, and are typical of the performance that can be obtained. The fixed point properties of projected cooling when $H(t)$ has only one bound state means that the method is flexible, efficient, and resistant to small errors. Given a Hamiltonian with a translationally-invariant kinetic energy term and interactions which go to 0 at large distances, projected cooling is able to construct a localized ground state efficiently and accurately.

5

Conclusions

In this thesis, we presented the theory used to develop two algorithms which have promising applications in nuclear physics. We showed these algorithms and presented significant results. The first algorithm we looked at is the trimmed sampling algorithm. We applied this algorithm to three different models and found strong results in each. In the case of the Bose-Hubbard model, we saw that while eigenvector continuation produces excellent results without noise, when we introduced noise to matrix elements by truncating the precision of the digits the accuracy of the results was no longer sufficient to reconstruct the behavior of the exact solution. We applied trimmed sampling with relatively simple filtering criteria and saw a substantial improvement in the quality of the results. In the next example, we looked at the Heisenberg chain model. In this case, noise was introduced as a small gaussian noise on each matrix element. Without altering our selection criteria, we again corrected the inaccuracy caused by this noise to a much greater degree than ridge regression, and also saw its ability to improve the accuracy of not only energy eigenvalue measurements but matrix elements as well. Finally, with some additional improvements we showed that on a large generalized eigenvalue problem produced using generator coordinate method on the Lipkin model, with error introduced through truncation, we again reconstructed the behavior of the energy and the matrix elements of operators.

Next, we looked at the projected cooling algorithm and tested it on three models. We saw that when calculating the ground state of these systems the projected cooling algorithm performed much better than other common methods. Projected cooling is a flexible and versatile method and is resistant to small errors, making it highly suitable to quantum computing.

This thesis presents these two algorithms with the intent that they will prove useful in future developments in the field of nuclear many-body physics and other applications where they may become useful.

References

[1] E. SCHRÖDINGER. **An Undulatory Theory of the Mechanics of Atoms and Molecules**. *Phys. Rev.*, **28**:1049–1070, Dec 1926. 3

[2] DILLON FRAME, RONGZHENG HE, ILSE IPSEN, DANIEL LEE, DEAN LEE, AND ERMAL RRAPAJ. **Eigenvector continuation with subspace learning**. *Phys. Rev. Lett.*, **121**(3):032501, 2018. 10, 26

[3] DILLON K. FRAME. *Ab Initio Simulations of Light Nuclear Systems Using Eigenvector Continuation and Auxiliary Field Monte Carlo*. PhD thesis, Michigan State University, 2019. 10

[4] AVIK SARKAR. *EIGENVECTOR CONTINUATION: CONVERGENCE AND EMULATORS*. PhD thesis, Michigan State University, 2022. 10

[5] AVIK SARKAR AND DEAN LEE. **Convergence of Eigenvector Continuation**. *Phys. Rev. Lett.*, **126**(3):032501, 2021. 10, 26

[6] H. F. TROTTER. **On the Product of Semi-Groups of Operators**. *Proceedings of the American Mathematical Society*, **10**(4):545–551, 1959. 13

[7] DAVID LAWRENCE HILL AND JOHN ARCHIBALD WHEELER. **Nuclear constitution and the interpretation of fission phenomena**. *Phys. Rev.*, **89**:1102–1145, 1953. 26

[8] JAMES J. GRIFFIN AND JOHN A. WHEELER. **Collective Motions in Nuclei by the Method of Generator Coordinates**. *Phys. Rev.*, **108**:311–327, 1957. 26

[9] CHUN WA WONG. **Generator-coordinate methods in nuclear physics**. *Physics Reports*, **15**(5):283–357, 1975. 26

[10] P. MAGIERSKI, P. H. HEENEN, AND W. NAZAREWICZ. **Generator-coordinate method study of hexadecapole correlations in superdeformed Hg-194**. *Phys. Rev. C*, **51**:R2880–R2884, 1995. 26

[11] K. VARGA AND Y. SUZUKI. **Precise Solution of Few Body Problems with Stochastic Variational Method on Correlated Gaussian Basis**. *Phys. Rev. C*, **52**:2885–2905, 1995. 26

[12] D. BLUME AND K. M. DAILY. **Universal relations for a trapped four-fermion system with arbitrary s -wave scattering length**. , **80**(5):053626, November 2009. 26

[13] TOMOAKI TOGASHI, YUSUKE TSUNODA, TAKAHARU OTSUKA, NORITAKA SHIMIZU, AND MICHIO HONMA. **Novel shape evolution in Sn isotopes from magic numbers 50 to 82**. *Phys. Rev. Lett.*, **121**(6):062501, 2018. 26

[14] P. DEMOL, T. DUGUET, A. EKSTRÖM, M. FROSINI, K. HEBELER, S. KÖNIG, D. LEE, A. SCHWENK, V. SOMÀ, AND A. TICHAI. **Improved many-body expansions from eigenvector continuation**. *Phys. Rev. C*, **101**(4):041302, 2020. 26

[15] S. KÖNIG, A. EKSTRÖM, K. HEBELER, D. LEE, AND A. SCHWENK. **Eigenvector Continuation as an Efficient and Accurate Emulator for Uncertainty Quantification**. *Phys. Lett. B*, **810**:135814, 2020. 26

[16] ANDREAS EKSTRÖM AND GAUTE HAGEN. **Global sensitivity analysis of bulk properties of an atomic nucleus**. *Phys. Rev. Lett.*, **123**(25):252501, 2019. 26

[17] R. J. FURNSTAHL, A. J. GARCIA, P. J. MILLICAN, AND XILIN ZHANG. **Efficient emulators for scattering using eigenvector continuation**. *Phys. Lett. B*, **809**:135719, 2020. 26

[18] ALFIO QUARTERONI, ANDREA MANZONI, AND FEDERICO NEGRI. *Reduced Basis Methods for Partial Differential Equations: An Introduction*, **92**. Springer, 2015. 26

[19] EDGARD BONILLA, PABLO GIULIANI, KYLE GODBEY, AND DEAN LEE. **Training and Projecting: A Reduced Basis Method Emulator for Many-Body Physics**. 3 2022. 26

[20] J. A. MELENDEZ, C. DRISCHLER, R. J. FURNSTAHL, A. J. GARCIA, AND XILIN ZHANG. **Model reduction methods for nuclear emulators**. *J. Phys. G*, **49**(10):102001, 2022. 26

[21] LIUMING LIU, GRAHAM MOIR, MICHAEL PEARDON, SINEAD M. RYAN, CHRISTOPHER E. THOMAS, POL VILASECA, JOZEF J. DUDEK, ROBERT G. EDWARDS, BALINT JOO, AND DAVID G. RICHARDS. **Excited and exotic charmonium spectroscopy from lattice QCD**. *JHEP*, **07**:126, 2012. 26

[22] ROBERT G. EDWARDS, NILMANI MATHUR, DAVID G. RICHARDS, AND STEPHEN J. WALLACE. **Flavor structure of the excited baryon spectra from lattice QCD**. *Phys. Rev. D*, **87**(5):054506, 2013. 26

[23] SERDAR ELHATISARI. **Adiabatic projection method with Euclidean time subspace projection**. *The European Physical Journal A*, **55**(144), 2019. 26

[24] SHIHANG SHEN, TIMO A. LÄHDE, DEAN LEE, AND ULF-G. MEISSNER. **Wigner SU(4) symmetry, clustering, and the spectrum of ^{12}C**. *Eur. Phys. J. A*, **57**(9):276, 2021. 26

[25] SHIHANG SHEN, TIMO A. LÄHDE, DEAN LEE, AND ULF-G. MEISSNER. **Emergent geometry and duality in the carbon nucleus**. 2 2022. 26

[26] AKHIL FRANCIS, ANJALI A. AGRAWAL, JACK H. HOWARD, EFEKAN KÖKCÜ, AND A. F. KEMPER. **Subspace Diagonalization on Quantum Computers using Eigenvector Continuation**. 9 2022. 27

[27] ALBERTO PERUZZO, JARROD MCCLEAN, PETER SHADBOLT, MAN-HONG YUNG, XIAO-QI ZHOU, PETER J. LOVE, ALÁN ASPURU-GUZIK, AND JEREMY L. O'BRIEN. **A variational eigenvalue solver on a photonic quantum processor**. *Nature Communications*, **5**:4213, July 2014. 27, 63

[28] E. F. DUMITRESCU, A. J. MCCASKEY, G. HAGEN, G. R. JANSEN, T. D. MORRIS, T. PAPENBROCK, R. C. POOSER, D. J. DEAN, AND P. LOUGOVSKI. **Cloud Quantum Computing of an Atomic Nucleus**. *Phys. Rev. Lett.*, **120**(21):210501, 2018. 27

[29] ZHENGRONG QIAN, JACOB WATKINS, GABRIEL GIVEN, JOEY BONITATI, KENNETH CHOI, AND DEAN LEE. **Demonstration of the Rodeo Algorithm on a Quantum Computer**. 10 2021. 27

[30] MAX BEE-LINDGREN, ZHENGRONG QIAN, MATTHEW DECROSS, NATALIE C. BROWN, CHRISTOPHER N. GILBRETH, JACOB WATKINS, XILIN ZHANG, AND DEAN LEE. **Rodeo Algorithm with Controlled Reversal Gates**. 8 2022. 27

[31] ANDREY NIKOLAYEVICH TIKHONOV. **On the stability of inverse problems**. *Doklady Akademii Nauk SSSR*, **39**:195–198, 1943. 27

[32] M. I. ALHEETY AND B. M. GOLAM KIBRIA. **Choosing Ridge parameters in the linear regression model with AR(1) error: A comparative simulation study**. *International Journal of Statistics & Economics*, **7**(A11):10–26, 2011. 27

[33] CALEB HICKS AND DEAN LEE. **Trimmed sampling algorithm for the noisy generalized eigenvalue problem**. *Phys. Rev. Res.*, **5**:L022001, Apr 2023. 27

[34] A. GELMAN, J.B. CARLIN, H.S. STERN, AND D.B. RUBIN. *Bayesian Data Analysis*. Chapman and Hall/CRC, 1995. 27

[35] D. R. PHILLIPS ET AL. **Get on the BAND Wagon: A Bayesian Framework for Quantifying Model Uncertainties in Nuclear Dynamics**. *J. Phys. G*, **48**(7):072001, 2021. 27

[36] W. K. HASTINGS. **Monte Carlo Sampling Methods Using Markov Chains and Their Applications**. *Biometrika*, **57**:97–109, 1970. 30

[37] A. F. M. SMITH AND A. E. GELFAND. **Bayesian Statistics without Tears: A Sampling–Resampling Perspective**. *The American Statistician*, **46**(2):84–88, 1992. 30

[38] WEIGUANG JIANG AND CHRISTIAN FORSSÉN. **Bayesian probability updates using sampling/importance resampling: Applications in nuclear theory**. *Front. in Phys.*, **10**:1058809, 2022. 30

[39] W. HEISENBERG. **Zur Theorie des Ferromagnetismus**. *Zeitschrift fur Physik*, **49**(9-10):619–636, September 1928. 37

[40] BING-NAN LU, NING LI, SERDAR ELHATISARI, DEAN LEE, EVGENY EPELBAUM, AND ULF-G. MEISSNER. **Essential elements for nuclear binding**. *Physics Letters B*, **797**:134863, 2019. 63

[41] SERDAR ELHATISARI, DEAN LEE, GAUTAM RUPAK, EVGENY EPELBAUM, HERMANN KREBS, TIMO A LÄHDE, THOMAS LUU, AND ULF-G. MEISSNER. **Ab initio alpha–alpha scattering**. *Nature*, **528**, 2015. 63

[42] ALEXEI Y. KITAEV. **Quantum measurements and the Abelian Stabilizer Problem**. *Electron. Colloquium Comput. Complex.*, **TR96**, 1995. 63

[43] DANIEL S. ABRAMS AND SETH LLOYD. **Simulation of Many-Body Fermi Systems on a Universal Quantum Computer**. *Phys. Rev. Lett.*, **79**:2586–2589, Sep 1997. 63

[44] E. F. DUMITRESCU, A. J. MCCASKEY, G. HAGEN, G. R. JANSEN, T. D. MORRIS, T. PAPENBROCK, R. C. POOSER, D. J. DEAN, AND P. LOUGOVSKI. **Cloud Quantum Computing of an Atomic Nucleus**. *Phys. Rev. Lett.*, **120**:210501, May 2018. 63

[45] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda. **A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem**. *Science*, **292**(5516):472–475, 2001. 63

[46] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. **Quantum Computation by Adiabatic Evolution**, 2000. 63

[47] David B. Kaplan, Natalie Klco, and Alessandro Roggero. **Ground States via Spectral Combing on a Quantum Computer**, 2017. 63

[48] Hefeng Wang. **Quantum algorithm for preparing the ground state of a system via resonance transition**, 2017. 63

[49] P. Oscar Boykin, Tal Mor, Vwani Roychowdhury, Farrokh Vatan, and Rutger Vrijen. **Algorithmic cooling and scalable NMR quantum computers**. *Proceedings of the National Academy of Sciences*, **99**(6):3388–3393, 2002. 63

[50] Jin-Shi Xu, Man-Hong Yung, Xiao-Ye Xu, Sergio Boixo, Zheng-Wei Zhou, Chuan-Feng Li, Alán Aspuru-Guzik, and Guang-Can Guo. **Demon-like algorithmic quantum cooling and its realization with quantum optics**. *Nature Photonics*, **8**(2):113–118, January 2014. 63

[51] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller. **Preparation of entangled states by quantum Markov processes**. *Phys. Rev. A*, **78**:042307, Oct 2008. 63

[52] Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac. **Quantum computation and quantum-state engineering driven by dissipation**. *Nature Physics*, **5**(9):633–636, September 2009. 63

[53] Dean Lee. **Lattice simulations for few- and many-body systems**. *Progress in Particle and Nuclear Physics*, **63**(1):117–154, 2009. 64

[54] Timo A Lähde and Ulf-G Meissner. *Nuclear lattice effective field theory: An introduction*, **957**. Springer, 2019. 64