SHELL MODEL THE MONTE CARLO WAY

W. E. Ormand

Proceedings of International Workshop on Structure and Dynamics of Quantum Many-Body Systems

Aizn, Japan

October 19-21, 1994
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
Shell Model the Monte Carlo Way

W. E. Ormand

Physics Department, 401 Nielsen Hall, University of Tennessee, Knoxville, TN 37996-1200

and

Oak Ridge National Laboratory, P.O. Box 2008, MS-6373 Building 6003, Oak Ridge, TN 37831-6373

Abstract

The formalism for the auxiliary-field Monte Carlo approach to the nuclear shell model is presented. The method is based on a linearization of the two-body part of the Hamiltonian in an imaginary-time propagator using the Hubbard-Stratonovich transformation. The foundation of the method, as applied to the nuclear many-body problem, is discussed. Topics presented in detail include: (1) the density-density formulation of the method, (2) computation of the overlaps, (3) the sign of the Monte Carlo weight function, (4) techniques for performing Monte Carlo sampling, and (5) the reconstruction of response functions from an imaginary-time auto-correlation function using MaxEnt techniques. Results obtained using schematic interactions, which have no sign problem, are presented to demonstrate the feasibility of the method, while an extrapolation method for realistic Hamiltonians is presented. In addition, applications at finite temperature are outlined.

1 Introduction

The shell model is one of the most successful descriptions of many-fermion systems including atomic nuclei [1]. In this picture, valence particles are spatially confined by a one-body potential and influence each other via a residual two-body interaction. Mathematically, the shell model can be reduced to a matrix-diagonalization problem by computing the matrix elements of the Hamiltonian between a set of basis states that span the configuration space of interest. Considerable effort has gone into studying nuclei within...
the framework of the shell model (developing effective interactions and operators etc.),
and impressive agreement between theoretical calculations and experimental data has
been achieved for nuclei with $A \leq 48$ [2, 3, 4].

The shell-model approach is limited primarily by the combinatorial growth in the
number of basis states with both the number of valence particles ($N_v$) and the size of
the single-particle basis ($N_s$). Indeed, for nuclei with $A \sim 60$, an unrestricted shell-
model calculation utilizing the $0f_{7/2}$-$0f_{5/2}$-$1p_{3/2}-1p_{1/2}$ orbits would involve approximately
$2 \times 10^9$ basis states with definite $z$-projection of angular momentum [5]. Although angular
momentum and isospin projection can reduce the size of the basis to about $10^7$, a problem
of this magnitude clearly lies beyond the capability of today’s computers.

The traditional approach to circumvent the computational limitations inherent in the
shell model is to impose what are often severe and ad hoc truncations on the number of
basis states. Unfortunately, because of the strong character of the residual interaction,
calculations of this nature can be unreliable, and significant renormalizations of the
residual interaction and transition operators are required.

In these lectures, a method that circumvents the “counting problem” inherent in the
matrix diagonalization approach will be discussed in detail. The algorithm is based on using
the imaginary-time propagator $\hat{U} = \exp(-\beta \hat{H})$ to either perform a thermodynamical
trace (canonical or grand canonical) at a temperature $T = \beta^{-1}$, or, for large $\beta$, to filter a
many-body trial state to the exact ground state. By applying the Hubbard-Stratonovich
transformation [6], the two-body terms in $\hat{U}$ are linearized with the introduction of an
auxiliary field, and the problem is transformed into a multi-dimensional integral that
must be evaluated using Monte Carlo techniques. The dimension of the integral, how-
ever, scales more gently with either $N_v$ or $N_s$. Hence, it will be shown that the method
that can give an exact treatment of a nuclear shell-model Hamiltonian, $\hat{H}$, even in situ-
ations where the matrix-diagonalization technique is impractical.

The outline of these lectures is as follows. In Section 2, a brief description of the
traditional, i.e. matrix-diagonalization, shell model is given. Presented in Section 3 is a
description of the formalism for the Monte Carlo approach to the shell model, including
a discussion of the sign of the Monte Carlo integrand. Section 4 contains a discussion re-
garding some computational details with the Monte Carlo shell model. Results comparing
the Monte Carlo approach with exact solutions obtained with the traditional approach
are shown in Section 5. In Section 6, techniques to treat realistic interactions are pre-
sented. Methods to obtain the strength function for transition operators are discussed in
Section 7. Applications regarding finite temperature are outlined in Section 8. Finally, concluding remarks are gathered in Section 9.

The reader should note that much of the material that I present here has been published previously [7, 8, 9, 10, 11, 12]. In addition, I wish to acknowledge collaborating on this project over the past three years with C. W. Johnson, G. H. Lang, S. E. Koonin, D. J. Dean, Y. Alhassid, P. B. Radha, and K. Langanke.

As a final note in this section, it should also be pointed out that the techniques outlined here have also been applied to problems in other fields in physics, and in particular the Hubbard model. I refer the reader to refs. [13, 14, 15, 16, 18] for the details of these applications.

2 Traditional Shell Model

In this section, I give a brief description of the standard approach to the shell-model problem in nuclear physics. Because of space limitations, it is necessary that I omit many of the details involved in performing shell-model calculations. Therefore, I refer the reader to refs. [19, 20, 21] for more details.

The goal of a nuclear structure theorist is to describe the properties of nuclei as accurately as possible. This begins with the nuclear Hamiltonian, which, if we restrict ourselves to two-body interactions, may be written as

\[ H = \sum_i t_i + \frac{1}{2} \sum_{ij} v_{ij}, \]  

where \( t_i \) and \( v_{ij} \) represent the kinetic and potential energy, respectively. Currently, exact solutions to the eigenvalue problem with the Hamiltonian Eq.(1) are possible only for \( A \leq 3 \) using the Faddeev equation [22]. In addition, a Green's function Monte Carlo approach has been used to obtain an "exact" solution for \(^4\text{He} \) [23]. For \( A \geq 5 \), however, there is no technique available that can give an exact solution to the many-body problem. Two approximate solutions are variational Monte Carlo methods [24] and the shell model as described here and in refs. [19, 20, 21].

The fundamental tenet of the shell model is to introduce a mean field potential \( U_i \), so that

\[ H = \sum_i (t_i + U_i) + \frac{1}{2} \sum_{ij} (v_{ij} - 2U_i\delta_{ij}) \]  

\[ = H_0 + H_{\text{res}}. \]
For the first part of Eq.(3), $H_0$, the eigenvalues can be obtained exactly, yielding a set of single-particle wave functions, which serve as a basis to construct a set of many-body Slater determinants to diagonalize the Hamiltonian $H$. In principle, the eigenvalues of $H$ can be obtained exactly provided all possible Slater determinants are accounted for in the diagonalization of the residual interaction. Choices of the mean field potential, $U$, can be the Hartree-Fock potential, a Woods-Saxon potential, or even a harmonic oscillator well.

The basic concept of the mean-field potential, $U$, is illustrated in Fig. 1, where the eigenstates of $U$, referred to single-particle states, are labeled on the right-hand side. In the shell model, one begins by occupying the lowest orbits, as indicated in the figure. The first thing to be noticed is the presence of energy gaps between groups of orbits. Notice also that these groups (or bundles) are generally of the same character, namely they have the same parity, and in terms of the harmonic oscillator, they usually have the same major oscillator quantum number. As such, the gaps lead to the notion of shell closures, in particular, the concept of a closed, inert core and an active valence space. This is illustrated in Figure 1 for 12 particles, where the first eight particles fully occupy the $0s$ and $0p$ orbits, and the remaining four valence particles are then permitted to occupy all possible states in the $0d$ and $1s$ orbits. Because of the computational limitations outlined below, shell-model calculations usually do not allow excitations of particles out of the core into the valence orbits. In addition, for the same reasons, the valence space is restricted to a few orbits, usually one major oscillator shell.

The procedure of the shell model can now be outlined as follows:

1. Obtain the valence single-particle basis in a given configuration space; e.g. $ls_{1/2}$, $0d_{5/2}$, $0d_{3/2}$ orbits

2. Construct the set of many-particle Slater determinants $|\psi_i\rangle$ spanning the many-body Hilbert space. Generally, the number of basis states is reduced by imposing symmetries such as the conservation of the third component of the angular momentum, or by projection, the total angular momentum and isospin;

3. Compute the Hamiltonian matrix $H_{ij} = \langle \psi_j | H | \psi_i \rangle$;

4. Diagonalize $H_{ij}$ to obtain the eigenvalues and eigenvectors. Usually, one is interested in only the lowest few eigenvalues, and an algorithm such as Lanczos [25] is used.
Figure 1: A schematic illustration of the single-particle potential $U_i$. The orbits are labeled on the right-hand side. As an illustration, twelve particles occupy the single-particle states. The first eight close the 0s and 0p orbitals, forming the inert closed core. The remaining four particles are then constrained to the 0d and 1s orbitals.

Several computer programs exist that implement the shell-model algorithm; each exploiting different features in order to achieve optimum performance for different aspects of the shell-model problem, e.g., binding energies, transition matrix elements, etc. The different programs, however, can be divided into two types: (1) those that build configuration basis states using $m$-scheme Slater determinants, or (2) those that use coefficients of fractional parentage (CFP). In the $m$-scheme, the Slater determinants have definite $z$-projection of the total angular momentum, $M$, and are represented as integer words on the computer, thereby exploiting the binary characteristics of Fermi occupation factors. One of the first codes to implement this approach was the Glasgow-Manchester Code [26]. Within this framework, angular momentum projection is achieved by noting
that the Hamiltonian matrix is a rotational invariant, and, therefore, the eigenvalues have definite angular momentum. One drawback of the \( m \)-scheme procedure is that after diagonalization, all states with \( J \geq M \) are obtained, and if states of only a particular spin and parity are desired, considerable effort is put into obtaining unwanted states. In addition, one is forced to diagonalize a much larger matrix than might be necessary, although progress has been made in improving the efficiency of the matrix diagonalization procedure with the shell-model code ANTOINE \cite{27}. The goal of the code OXBASH \cite{28} is to reduce the dimensions of the matrix to be diagonalized by projecting angular momentum onto the \( m \)-scheme basis, and, therefore, focusing on states with definite angular momentum (and isospin if desired) and parity.

Instead of using \( m \)-scheme Slater determinants, it is also possible to construct anti-symmetric, angular-momentum projected basis states with definite angular momentum (and isospin) using the technique of CFP's. The basis states are constructed by adding one particle to the anti-symmetrized \( n - 1 \) states via

\[
\psi(j^n, \alpha JM) = \sum_{\beta, J'} [j^{n-1}(\beta, J')j | j^n \alpha J] [\psi(j^{n-1}, \beta J') \otimes j]^{JM},
\]

where \( \psi(j^{n-1}, \beta J') \) represents the anti-symmetrized basis states with \( n - 1 \) particles and \( [j^{n-1}(\beta, J')j | j^n \alpha J] \) is the coefficient of fractional parentage. One of the first computer codes to implement the CFP procedure is the Oak Ridge Code \cite{29}. Improvements to the general procedure using permutation group symmetries are now being implemented in the Drexel University Shell Model (DUSM) code \cite{30}.

The primary reason for restricting the shell-model problem to just the valence particles occupying a few orbits is to limit the number of Slater determinants to a manageable size so that the diagonalization of the Hamiltonian is computationally feasible. The total number of Slater determinants within a Hilbert space is given by the relation

\[
\binom{N_p^n}{N_p^2} \binom{N_n^n}{N_n^2} = \frac{N_p^n!}{(N_p^2 - N_p^n)! N_p^2! (N_p^n - N_n^n)! N_n^n!},
\]

where \( N_p^{(n)} \) and \( N_n^{(n)} \) are the number of proton (neutron) single-particle states and valence particles, respectively. In actuality, Eq.(5) overestimates the dimension of the basis states since we may also fix the third component of angular momentum (\( m \)-scheme) or even project total angular momentum (\( J \)-projection). Nonetheless, Eq.(5) gives a rough estimate of the number of \( m \)-scheme Slater determinants to within a factor of 10-100. For example, for an \( fp \)-shell calculation of \( ^{60}\text{Zn} \), with \( N_p = 20 \) and \( N_n = 10 \) for
both protons and neutrons, Eq.(5) gives an estimate of $3.4 \times 10^{10}$ states, whereas the number $m$-scheme states with \( J_z = 0 \) is \( \approx 2 \times 10^9 \). On the other hand, by projecting angular momentum, the number of states with \( J = 0 \) is of the order \( 10^7 \). For the most part, a calculation of this magnitude is beyond the capabilities of any computer currently available. In particular, the storage requirements for this case would easily exceed 100 Gbytes. Currently, a typical large-basis shell-model calculation that will run overnight on most workstations has dimensions of the order \( 10^5-6 \) Slater determinants with total \( J_z \), or of the order 5000-10000 states with projected angular momentum and isospin.

From Eq.(5), it is clear that the matrix-diagonalization approach to the shell model will always be computationally limited by the number of states that constitute the basis. This is true even when we limit ourselves to a computation using only one major oscillator shell, as is illustrated in the $^{60}\text{Zn}$ example above. As such, the matrix-diagonalization approach requires a truncation on the number of basis states, even within one major oscillator shell. One potentially dangerous aspect of truncating within a major oscillator shell is that for the most part the nucleon-nucleon interaction commutes with the generators of the SU(4) group [31]. The primary component of the nuclear Hamiltonian that breaks SU(4) symmetry is the spin-orbit splitting of the single-particle energies. Hence, if one starts with basis states composed of the lowest orbitals, SU(4) is maximally violated. However, the effect of including all configurations within the major shell is to restore SU(4) symmetry to the level of approximately 50% [32]. This partial restoration of SU(4) symmetry can have important consequences regarding the strength of transition operators. In particular, in the SU(4) limit, Gamow-Teller transitions can occur only between states in the same SU(4) multiplet. The partial restoration of SU(4) symmetry within a full space shell-model calculation then is an possible explanation for the considerable decrease of the Gamow-Teller strength (B(GT) values) from the single-particle estimate when the full space is considered. Because of this, one might conclude that full-space calculations within at least one major shell are probably crucial for the prediction of some quantities.

Finally, in regards to solving a shell-model problem for a realistic description of nuclei, there are essentially two distinct problems that must be faced. The first is the choice of the shell-model Hamiltonian itself. Because of computational limitations, shell-model calculations are generally restricted to excitations within a reduced valence-particle space. The validity of such a limitation can, of course, be estimated with perturbation theory,
and it is generally found that with realistic nucleon-nucleon interactions, excitations of the core can play a significant role. For this reason, shell-model calculations are performed with effective interactions that are designed to incorporate excitations that are omitted from the configuration space. One attempt to accomplish this is the $G$-matrix, as was done for the $fp$ shell by Kuo and Brown [33]. Even here, though, agreement with experimental data is not as good as one might have hoped, indicating that further renormalization is required [34]. For the most part, this additional renormalization is accomplished empirically by fitting components of the Hamiltonian to experimental data. An excellent example is the $sd$-shell Hamiltonian of Wildenthal [2], where several hundred levels are reproduced to within a few hundred keV. The second problem associated with shell-model calculations, is obtaining an exact solution for a given Hamiltonian within a given model space. Primarily, this is a computational problem, as the principal limitation is accounting for all the excitations that lie within the configuration space. In general, I feel that the matrix-diagonalization approach is the method of choice wherever it is computationally feasible, as it can lead to a wealth of information not available in the method presented in this lecture. Nonetheless, this approach will always reach a point at which it is computationally not practical to use. Because of the rapid growth in the basis dimensions, this limitation can be quickly achieved with the addition of one or two particles within a large configuration space. Hence, even with the rather heroic efforts of the Strasbourg group to complete full-space calculations in the $fp$-shell for $A=50$ using the code ANTOINE [35], it is unlikely that full-space calculations can be carried out for $A=60$. In addition, for nuclei with $A\approx70$, the $0g_{9/2}$ orbit becomes important, and as is apparent from Eq.(5), the dimensions required for these nuclei will be larger even still. Given that full-space calculations might be required to adequately describe some features of nuclear structure, such as Gamow-Teller transitions, it appears as if theoretical approaches to detailed nuclear structure are doomed to failure because of an exploding number of configurations and not enough computational resources. In these lectures, however, I will describe a new technique that is able to circumvent the “counting problem” inherent in shell-model calculations, and in many cases give an “exact” solution even though the traditional matrix-diagonalization approach is hopeless.
3 Monte Carlo Approach

In this section, a full description of the formalism required to perform basic calculations for the nuclear shell-model problem using auxiliary-field Monte Carlo (AFMC) techniques is presented. This section, being the longest in these notes, is organized in the following manner. In subsection 3.1, the Monte Carlo approach is derived for the general Hamiltonian problem using the Hubbard-Stratonovich transformation. In subsection 3.2 a procedure to bring the nuclear Hamiltonian into quadratic form is outlined. The mathematical formulae associated with computing the overlaps are given in subsection 3.3, and a discussion regarding the sign of the Monte Carlo integrand is given in subsection 3.4. Approximate solutions to the Monte Carlo integrand are discussed in subsection 3.5, and finally, a techniques for performing the Monte Carlo integration are presented in subsection 3.6.

3.1 The Hubbard-Stratonovich Transformation

The Monte Carlo approach to the shell model is motivated by the observation that the expectation value of any operator, $\hat{O}$, in the ground state of some Hamiltonian, $\hat{H}$, can be obtained via

$$\langle \hat{O} \rangle_{GS} = \lim_{\beta \to \infty} \frac{\langle \psi_t | e^{-\beta \hat{H}/2} \hat{O} e^{-\beta \hat{H}/2} | \psi_t \rangle}{\langle \psi_t | e^{-\beta \hat{H}} | \psi_t \rangle},$$

where $\psi_t$ is any trial wave function not orthogonal to the ground state. In addition, there are times when we are interested in the properties of nuclei at finite excitation energy. This can be achieved by the introduction of a Lagrange multiplier, $\beta$, to fix the average due of the expectation value of the $\hat{H}$, as is done in statistical mechanics. Namely we are interested in the partition function

$$Z = \hat{\text{Tr}} \exp(-\beta \hat{H}).$$

Here, $\beta$ may be interpreted as an inverse temperature ($T = 1/\beta$), and $\hat{\text{Tr}}$ represents either a grand-canonical trace over all many-body states in the space (denoted by $\hat{\text{Tr}}_G$) or a canonical trace with fixed particle number $N_o$ (denoted by $\hat{\text{Tr}}_N_o$). Given the partition function in Eq. (7), the thermal observable of an operator $\hat{O}$ is

$$\langle \hat{O} \rangle = \frac{1}{Z} \hat{\text{Tr}} [\hat{O} \exp(-\beta \hat{H})].$$

Note that the operator $\hat{\mathcal{U}} = e^{-\beta \hat{H}}$ is similar in form to a time propagator in quantum mechanics except that the exponent is purely real. On the other hand, if we define $\beta = it$, 

$$\hat{\mathcal{U}} = e^{-it\hat{H}},$$

then the propagator is a complex function of $t$. However, in the case of imaginary time $\beta = it$, the propagator reduces to a real function. The imaginary time formalism is particularly useful in the context of quantum mechanics, as it allows for the use of the path integral formulation of quantum field theory. The path integral formalism provides a powerful tool for calculating expectation values and other observables of quantum systems, offering a more intuitive and conceptually cleaner approach compared to the direct calculation of matrix elements. In the context of quantum field theory, the path integral over all possible histories of a field is used to compute physical observables, with the Gaussian integral over all intermediate states playing a crucial role in simplifying the calculations. In the case of quantum mechanics, the path integral representation of the propagator allows for a more straightforward computation of the propagation of a quantum state through a potential, providing a direct connection to the classical mechanics of the system.
then we may think of $\hat{U}$ as an “imaginary”-time propagator, and I will often use this terminology to describe $\hat{U}$.

In what follows, the applicable formulae are presented using the notation associated with the finite-temperature formalism. However, the subsequent derivations are also valid for the zero-temperature formalism of Eq.(6) by making the substitution

$$\text{Tr}(\ ) \rightarrow \langle \psi_t \mid \psi_t \rangle.$$  \hfill (9)

For the Monte Carlo approach to the shell model, we restrict ourselves to Hamiltonians that contain at most two-body terms, such that in second-quantization formalism

$$\hat{H} = \sum_\alpha \epsilon_\alpha a_\alpha^\dagger a_\alpha + \frac{1}{2} \sum_\alpha \beta \gamma \delta \epsilon_\alpha a_\alpha^\dagger a_\beta^\dagger a_\gamma a_\delta,$$  \hfill (10)

where $a_\alpha^\dagger$ and $a_\alpha$ are the anti-commuting creation and annihilation operators associated with the single-particle state $\alpha$ defined by the complete set of quantum numbers $nljmmt$, ($n$, $l$, $j$, $m$, and $t_z$ denote the principal, orbital angular momentum, total single-particle angular momentum, $z$-projection of $j$, and the third component of the isospin quantum numbers, respectively), and $\epsilon_\alpha$ and $V_{\alpha\beta\gamma\delta}$ are the single-particle energies and two-body matrix elements of the residual interaction. Note that for any two-body Hamiltonian of the form Eq. (10), it is possible to find a set of one-body operators $\hat{O}_\alpha$ so that $\hat{H}$ may be written as

$$\hat{H} = \sum_\alpha \epsilon_\alpha \hat{O}_\alpha + \frac{1}{2} \sum_\alpha V_\alpha \hat{O}_\alpha^2.$$  \hfill (11)

At this stage, neither Eq.(6) or Eq.(8) lead to an improvement over the matrix diagonalization approach. This is because of the two-body character of $\hat{H}$ in $\hat{U}$, which, when applied to any trial wave function, will span into the complete Hilbert space. It would, in fact, be much more convenient if instead $\hat{U}$ was the exponential of a one-body operator since the application of $\hat{U}$ on a trial wave function that is a Slater determinant would return, by Thouless’ theorem [36], another single Slater determinant.

To achieve a simplification of the imaginary-time propagator $\hat{U}$, we make use of the Hubbard-Stratonovich transformation [6], which is based on the identity

$$e^{\frac{1}{2} \Lambda \hat{O}^2} = \sqrt{\frac{|\Lambda|}{2\pi}} \int \mathrm{d} \sigma e^{-\frac{1}{2} |\Lambda| \sigma^2 + \sigma \Lambda \hat{O}},$$  \hfill (12)

where $\sigma = \pm 1$ if $\Lambda \geq 0$ or $\pm i$ if $\Lambda < 0$. We can then transform $\hat{U}$ into a one-body operator with the introduction of an auxiliary field $\sigma_\alpha$ for each operator $\hat{O}_\alpha$ via Eq.(12), giving

$$e^{-\beta \hat{H}} \approx \int \mathcal{D}(\sigma) \exp \left[-\frac{1}{2} \beta \sum_\alpha \sigma_\alpha^2 \right] e^{-\beta \mathcal{H}(\sigma)},$$  \hfill (13)
\[ D(\sigma) = \prod_{\alpha} \sqrt{\frac{\beta |V_{\alpha}|}{2\pi}} d\sigma_{\alpha} \]  \hspace{1cm} (14)

and

\[ \hat{h}(\sigma) = \sum_{\alpha} (\varepsilon_{\alpha} + s_{\alpha} V_{\alpha} \sigma_{\alpha}) \hat{\mathcal{O}}_{\alpha}. \]  \hspace{1cm} (15)

Notice the use of the \( \approx \) sign in Eq.(13). This is because although Eq.(12) is an exact identity, the operators \( \hat{\mathcal{O}}_{\alpha} \) do not necessarily commute, and when Eq. (12) is applied to \( \hat{U} \), the resulting integral is accurate only to first order in an expansion of \( e^{-\beta \hat{h}} \). The accuracy is improved by dividing \( \hat{U} \) into \( N_t \) “time slices”, so that

\[ \hat{U} = e^{-\frac{\Delta \beta \hat{H}}{N_t}} \ldots e^{-\frac{\Delta \beta \hat{H}}{N_t}} \]  \hspace{1cm} (16)

and applying Eq. (12) to each “slice”. For what follows, we introduce \( \tau_n \) as an imaginary time in the range \((0, \beta)\) defined as \( \tau_n = n\Delta \beta \), with \( \Delta \beta = \beta/N_t \).

The thermal observable in Eq. (8) can now be written as

\[ \langle \hat{O} \rangle = \frac{\int D[\sigma] G(\sigma) \langle \hat{O} \rangle_{\sigma} \zeta(\sigma)}{\int D[\sigma] G(\sigma) \zeta(\sigma)}, \]  \hspace{1cm} (17)

where \( D(\sigma) = \prod_{\alpha,n} \sqrt{\Delta \beta |V_{\alpha}| / 2\pi d\sigma_{\alpha}(\tau_n)} \), \( G(\sigma) = \exp[-\frac{1}{2} \Delta \beta \sum_{\alpha,n} |V_{\alpha}| \sigma_{\alpha}^2(\tau_n)] \), the overlap trace over the one-body, imaginary-time evolution operator is

\[ \zeta(\sigma) = \text{Tr} \left[ \hat{U}_{\sigma}(\beta, 0) \right] \]  \hspace{1cm} (18)

and

\[ \langle \hat{O} \rangle_{\sigma} = \frac{\text{Tr} \left[ \hat{O} \hat{U}_{\sigma}(\beta, 0) \right]}{\text{Tr} \left[ \hat{U}_{\sigma}(\beta, 0) \right]} \]  \hspace{1cm} (19)

In these expressions, the one-body evolution operator is defined as

\[ \hat{U}_{\sigma}(\tau_j, \tau_i) = \hat{U}(\sigma(\tau_j)) \ldots \hat{U}(\sigma(\tau_{i+1})) \hat{U}(\sigma(\tau_i)), \]  \hspace{1cm} (20)

with

\[ \hat{U}(\sigma) = \exp \left[ -\Delta \beta \hat{h}(\sigma) \right]. \]  \hspace{1cm} (21)

Equation (17) expresses the expectation value of any observable as a multi-dimensional integral whose dimensions are of the order \( N_t^2 \cdot N_t \), so that it must be evaluated by Monte
Carlo techniques. In order to perform a Monte Carlo integration, it is necessary to define a positive-definite weight function $W(\sigma)$. For the weight function, it is convenient to choose $W(\sigma) = \mathcal{G}(\sigma)|\zeta(\sigma)|$ so that

$$\langle \hat{\mathcal{O}} \rangle = \frac{\int \mathcal{D}[\sigma]W(\sigma)\langle \hat{\mathcal{O}} \rangle_\sigma \Phi(\sigma)}{\int \mathcal{D}[\sigma]W(\sigma)\Phi(\sigma)},$$

with the "sign" $\Phi(\sigma) \equiv \zeta(\sigma)/|\zeta(\sigma)|$. The observable $\langle \hat{\mathcal{O}} \rangle$ is then computed by selecting an ensemble $\{\sigma_k\}$ chosen according to distribution $W(\sigma)$, and computing the ensemble average; i.e.,

$$\langle \hat{\mathcal{O}} \rangle_{MC} = \frac{\sum_k \langle \hat{\mathcal{O}} \rangle_{\sigma_k} \Phi(\sigma_k)}{\sum_k \Phi(\sigma_k)},$$

The uncertainty in the Monte Carlo result is obtained by noting that if the "sign" and the observable are computed with the same set of fields, we must also account for correlations between the two observations, namely the uncertainty in the observable is

$$\delta\langle \hat{\mathcal{O}} \rangle_{MC} = \frac{1}{\sqrt{N}} \left[ \frac{\langle \hat{\mathcal{O}}^2 \Phi^2 \rangle - \langle \hat{\mathcal{O}} \Phi \rangle^2}{\langle \Phi \rangle^2} + \frac{\langle \hat{\mathcal{O}} \Phi \rangle^2}{\langle \Phi \rangle^4} \left( \langle \Phi^2 \rangle - \langle \Phi \rangle^2 \right) \right]^{1/2},$$

where $\langle \cdots \rangle$ denotes $\sum_k$. Now, note that if $\Phi(\sigma_k)$ oscillates wildly from sample to sample, large errors can occur in $\langle \hat{\mathcal{O}} \rangle$ because of the poorly determined numerator and denominator in Eq. (23).

Some remarks about the nature of the operators in Eq. (17) and their notation are now in order. First, since $\hat{h}(\sigma)$ is only a one-body operator, the evolution operator $\hat{U}(\sigma)$ can be represented as a $N_\nu \times N_\nu$ matrix $U(\sigma)$ in the single-particle basis: $U(\sigma) = \exp(-\Delta \mathbf{h}(\sigma))$, where $(\mathbf{h}(\sigma))_{\alpha\beta} = \langle \alpha | \hat{h}(\sigma) | \beta \rangle$. Likewise, any Slater determinant $|\psi_S\rangle$ describing $N_\nu$ particles can be represented by a matrix, $\Psi_S$, with $N_\nu$ columns and $N_\nu$ rows. Thouless' theorem then implies that the operation $\hat{U}(\sigma)|\psi_S\rangle$ yields only a single Slater determinant, thus averting the need to have all Slater determinants stored in order to evaluate Eq. (8). In what follows, we represent the product of one-body evolution operators $\hat{U}_\sigma(\tau_j, \tau_i)$ as

$$U_\sigma(\tau_j, \tau_i) = U(\sigma(\tau_j)) \cdots U(\sigma(\tau_{i+1}))U(\sigma(\tau_i)),$$

with the implicit understanding that $U_\sigma(\beta, 0) = U_\sigma$.

At this point, we now turn to a discussion on bringing the Hamiltonian into quadratic form.
3.2 Bringing the Hamiltonian into Quadratic Form

Since we have restricted ourselves to Hamiltonians that are at most two-body in nature, it is always possible to find a canonical transformation so that the Hamiltonian may be brought into the quadratic form of Eq.(11). What is more, the transformation is not unique and the choice may have important consequences regarding the sign of the Monte Carlo integral.

We begin by writing the Hamiltonian in terms of the $jj$-coupled two-body matrix elements, namely

$$\hat{H} = \sum_i e_i a_i^\dagger a_i$$

$$-\frac{1}{4} \sum_{ijkl, JT} V_{ij}^{\pm k l}(ij, k l) [(2J + 1)(2T + 1)(1 + \delta_{ij})(1 + \delta_{kl})]^{1/2}$$

$$[(a_i^\dagger \otimes a_j^\dagger)^{JT} \otimes (\bar{a}_k \otimes \bar{a}_l)^{JT}]^{00,00},$$

where $V_{ij}^{\pm k l}(ij, k l)$ are the anti-symmetrized, $jj$-coupled two-body matrix elements of the Hamiltonian. The choice of notation used in this section will be to use the letters $i, j, k, l$ to denote the single-particle orbits, i.e., $nlj$, whereas the Greek letters $\alpha, \beta, \gamma, \delta$ will be used to denote the explicit single-particle state, including the $j_z$ quantum number. At this point, we have two choices in which to bring the Hamiltonian into quadratic form. The first is in terms of the paring operators $A_{ij}^{+ T(ij)} = (a_i^\dagger \otimes a_j^\dagger)^{JT}$, while the second is a density-density decomposition with the operators $\hat{\rho}_{ij}^{KM} = (a_i^\dagger \otimes \bar{a}_j)^{KM}$. Although both formalisms are possible, it turns out that the density-density break-up is more convenient in terms of the sign of the Monte Carlo integral, even for a pure pairing interaction. For this reason, in these lecture I will focus on the density-density decomposition.

In order to illustrate the density-density break-up, we first consider the case of like particles, and ignore the isospin label in Eq.(26). We now make use of the commutation relation

$$a_i^\dagger a_j^\dagger a_\gamma a_\delta = a_i^\dagger a_\delta a_\gamma - a_i^\dagger a_\gamma a_\delta a_\delta$$

and recouple the angular momenta, to form a scalar operator. The Hamiltonian may now be written as

$$\hat{H} = \sum_{ij} \tilde{e}_{ij} \sqrt{2j_i} + 1 \rho_{ij}^{00} + \frac{1}{2} \sum_{ijkl, KM} E_{K}(ik, jl)(-1)^M \rho_{ik}^{KM} \rho_{jl}^{-M},$$

(28)
where

\[ \varepsilon_{ij} = \epsilon_i \delta_{ij} - \frac{1}{2} \sum_k \sum_J (-1)^{j_i+j_k} \frac{2J+1}{2j_i+1} \left( \frac{(1+\delta_{ij})(1+\delta_{kl})}{4} \right)^{1/2} \]  

(29)

and

\[ E_K(i,k,j,l) = \sum_j (-1)^{j_i+j_k+J} \left\{ \begin{array}{ccc} j_i & j_j & K \\ j_i & j_k & J \end{array} \right\} (2J+1) \left( \frac{(1+\delta_{ij})(1+\delta_{kl})}{4} \right)^{1/2} V_j^k(ij, kl). \]  

(30)

Now, we introduce the set of operators

\[ \hat{Q}_{ij}^{KM} = \begin{cases} \frac{1}{2} (\rho_{ij}^{K0} + (\rho_{ij}^{K0})^\dagger) & \text{if } M = 0 \\ \frac{1}{\sqrt{2}} (\rho_{ij}^{KM} + (\rho_{ij}^{KM})^\dagger) & \text{if } M \neq 0 \end{cases} \]  

(31)

\[ \hat{P}_{ij}^{KM} = \begin{cases} -\frac{i}{2} (\rho_{ij}^{K0} - (\rho_{ij}^{K0})^\dagger) & \text{if } M = 0 \\ -\frac{i}{\sqrt{2}} (\rho_{ij}^{KM} - (\rho_{ij}^{KM})^\dagger) & \text{if } M \neq 0 \end{cases} \]  

(32)

so that

\[ \hat{H} = \sum_{ij} \varepsilon_{ij} \sqrt{2j_i + 1} \rho_{ij}^{oo} + \frac{1}{2} \sum_{ijkl} E_K(i,k,j,l) \left[ \hat{Q}_{ij}^{KM} \hat{Q}_{jl}^{KM} + \hat{P}_{ij}^{KM} \hat{P}_{jl}^{KM} \right]. \]  

(33)

The final step towards bringing \( \hat{H} \) into quadratic form is to diagonalize the matrix \( E_K(i,k,j,l) \), leading to

\[ \hat{H} = \sum_{ij} \varepsilon_{ij} \sqrt{2j_i + 1} \rho_{ij}^{oo} + \frac{1}{2} \sum_{K,M} V_j^k \left[ (\hat{Q}_a^{KM})^2 + (\hat{P}_a^{KM})^2 \right]. \]  

(34)

Equation (34) is now in quadratic form and we can apply the Hubbard-Stratonovich transformation to the exponentiated form \( e^{-\Delta \hat{H}} \). Note that there are now auxiliary fields associated with each operator \( \hat{Q}_a^{KM} \) and \( \hat{P}_a^{KM} \), which for convenience we refer to as \( \sigma \) and \( \tau \) fields. In general, the maximum number of auxiliary fields is given by \( N_s^2 \) (\( N_s \) = the number of single-particle states) for each time slice.

At this point, we note that the Hamiltonian is completely specified by only the anti-symmetric two-body matrix elements of Eq.(26). However, we can add to the \( V_j^k(ij, kl) \) any set of symmetrized two-body matrix elements satisfying

\[ V_j^S(ij, kl) = -(\varepsilon)^{j_i+j_k-J-T} V_j^S(ij, kl). \]  

(35)

That such a freedom exists is due to the Pauli principle, since, for example, we have

\[ [a_i^+ \otimes a_i^+]_{J=1}^{M,T=1} | \psi \rangle = 0. \]  

(36)
As such, the added symmetrized matrix elements do not contribute to the expectation value of the Hamiltonian. They can, however, have an important impact on the form of the Hamiltonian after the density-density decomposition. An example is the case of the separable quadrupole-quadrupole (QQ) interaction, which by construction is of a density-density type, and we should be able to rewrite \( \hat{H} \) in terms of density operators with just \( K = 2 \). This is true for the direct part of the QQ interaction. The exchange part, however, transforms to all possible \( K \)-values, and no simplification is possible. As we will see, this "bad" behaviour of the exchange term can be compensated for by adding a set of symmetric two-body matrix elements. In what follows, we use a set of matrix elements that possess no defined symmetries to define the hamiltonian in Eq.(26), namely

\[
V_{ij kl}^N = V_{ij kl}^A + V_{ij kl}^S.
\]  

(37)

In regards to protons and neutrons, note that we can also re-couple the isospin quantum numbers after the density-density decomposition in Eq.(28); this yields the operators of the form \( \rho_{ij}^{K M, T T} \). One drawback of this breakup, however, is that the operators \( \rho_{ij}^{K M, T = \pm 1, T_3 = \pm 1} \) do not conserve the number of protons and neutrons in the application of \( e^{-\beta \hat{h}(\sigma)} \). As such, the matrix representing \( \hat{h}(\sigma) \) has dimensions \( (N_p + N_n) \times (N_p + N_n) \) and the proton and neutron number must be projected with a two-dimensional integral (see below). To some degree, it is more convenient to work with operators that explicitly conserve the number of protons and neutrons. The principle advantages of this choice are that \( \hat{h} \) is separated into two smaller parts and that the projection of proton and neutron number in the canonical ensemble is performed separately with two, one-dimensional sums.

We now perform the density-density decomposition using the operators

\[
\rho_{ij}^{K M, t} = \rho_{ij}^{K M, p} + (-1)^t \rho_{ij}^{K M, n},
\]  

(38)

where \( t = 0, 1 \) and the superscript \( p \) and \( n \) denote the proton and neutron densities, respectively. Now, in terms of both protons and neutrons, the Hamiltonian can be written as

\[
\hat{H} = \sum_{ij} \varepsilon_{ij} \sqrt{2j_i + 1} \rho_{ij}^{00, t=0} + \frac{1}{2} \sum_{ijkl} E_{K,t}(ik, jl)(-1)^M \rho_{ik}^{K M, t} \rho_{jl}^{K - M, t},
\]  

(39)

where

\[
E_{K,t=0}(ik, jl) = \sum_J (-1)^{j_i + j_k + J}(2J + 1) \left\{ \begin{array}{ccc} j_i & j_i & K \\ j_l & j_k & J \end{array} \right\} \sqrt{(1 + \delta_{ij})(1 + \delta_{kl})} \]

\[
\frac{1}{2} \left[ V_{ij kl}^N + \frac{1}{2} (V_{ij kl}^A + V_{ij kl}^S) \right]
\]  

(40)
\[ E_{K,t=1}(ik, jl) = \sum_{j} (-1)^{j_j+j_k+j} (2J+1) \left\{ \begin{array}{ccc} j_l & j_j & K \\ j_l & j_k & J \end{array} \right\} \sqrt{(1+\delta_{ij})(1+\delta_{kl})} \]

\[ \frac{1}{4} \left( V^A_{J,T=1}(ij, kl) - V^S_{J,T=1}(ij, kl) \right), \]  \hspace{1cm} (41) \]

with

\[ V^{S/A}_{J,T}(ij, kl) = \frac{1}{2} \left[ V^{N}_{J,T}(ij, kl) \pm (-1)^{j_j+j_k+j} V^{N}_{J,T}(ji, kl) \right]. \]  \hspace{1cm} (42) \]

Note that since \( V^{N}_{J,T}(ij, kl) \) is arbitrary, we can choose \( V^{N}_{J,T}(ij, kl) = V^{N}_{J,T=0}(ij, kl) \) so that \( E_{K,t=1}(ik, jl) = 0 \). The advantage of this choice is to reduce the number of density operators, and, therefore, the number of auxiliary fields. In addition, with this choice, the exchange term for pure density-density operators is cancelled, thereby reducing the number of auxiliary fields yet further.

Finally, the one-body Hamiltonian \( \hat{h}(\sigma) \) is given by

\[ \hat{h}(\sigma) = \sum_{ij} \hat{e}_{ij} \sqrt{2j_i + 1} \rho_{ij}^{00, t=0} + \sum_{K,M} \sum_{s=0}^{\infty} \alpha_{K,M} \sigma_{K,M}^{t=0} \left( \hat{Q}^{K,M,t} + \hat{\tau}^{K,M,t} \right), \]  \hspace{1cm} (43) \]

where \( \hat{Q}^{K,M,t} \) and \( \hat{\tau}^{K,M,t} \) are given by Eqs.(31,32) using the definition of \( \rho_{ij}^{K,M,t} \) of Eq.(38). Note at this point, that the basic building block of the matrix \( (h(\sigma))_{ij} \) is the matrix element of the density operator \( \rho_{ij}^{K,M} \), which is given by

\[ \langle j_\alpha m_\alpha | \rho_{ij}^{K,M} | j_\beta m_\beta \rangle = (-1)^{i_\beta-m_\beta} (j_\alpha m_\alpha j_\beta - m_\beta | K M) \delta_{i\alpha} \delta_{j\beta} \]  \hspace{1cm} (44) \]

### 3.3 Overlaps

In this section, the explicit form of the overlap and the density matrix are given. We begin with the zero-temperature formalism, where the Slater determinants are represented as a matrix with \( N_s \) columns and \( N_s \) rows. The required overlap is then

\[ \langle \psi_\epsilon | \hat{U}(\beta, 0) | \psi_\epsilon \rangle = \langle \psi_L(\tau) | \psi_R(\tau) \rangle = \det \left[ \Psi_L^\dagger \Psi_R \right], \]  \hspace{1cm} (45) \]

where

\[ \Psi_R \equiv U_\epsilon(\tau, 0) \Psi_\epsilon; \quad \Psi_L \equiv U_\epsilon^\dagger(\beta, \tau) \Psi_\epsilon \]  \hspace{1cm} (46) \]

are the matrices representing \( \psi_R \) and \( \psi_L \), respectively. By using Wick's theorem [37], the expectation value of any \( n \)-body operator can be can be expressed as a sum of products of expectation values of one-body operators. Therefore, the basic quantity of interest is

\[ \langle a_\alpha^\dagger a_\beta \rangle = \left[ \Psi_R \left( \Psi_L^\dagger \Psi_R \right)^{-1} \Psi_L^\dagger \right]_{\alpha \beta}. \]  \hspace{1cm} (47) \]
In particular, for a two-body operator, we have

\[ (a_d^\dagger a_b^\dagger a_d a_b)_\sigma = \langle a_d^\dagger a_b^\dagger \rangle_\sigma \langle a_d a_b \rangle_\sigma + \langle a_d^\dagger a_d \rangle_\sigma \langle a_b^\dagger a_b \rangle_\sigma. \]  

(48)

Given the matrix notation for \( \hat{U}_\sigma(\beta, 0) \), the grand-canonical trace is given by

\[ \text{Tr}_{G}[\hat{U}_\sigma(\beta, 0)] = \det(1 + U_\sigma), \]  

(49)

while the density matrix is given by

\[ \langle a_d^\dagger a_b \rangle_\sigma = [(1 + U_\sigma(\beta, 0))^{-1} U_\sigma(\beta, 0)]_{ba}. \]  

(50)

Again, the expectation value of two-body operators can be computed using Wick's theorem and Eq.(48).

For the canonical trace for fixed particle number \( N_\nu \), \( \zeta_{N_\nu}(\sigma) \), can be obtained from an activity expansion by noting that [8]

\[ \det(1 + \lambda U_\sigma) = \sum_{N=0}^{N_\nu} \lambda^N \zeta_N(\sigma) \]

\[ = \exp[\text{tr} \ln(1 + \lambda U_\sigma)] = \exp \left( \sum_{n=1}^{N_\nu} \frac{(-1)^{n-1}}{n} \lambda^n \text{tr}[U_\sigma^n] \right), \]  

(51)

where the symbol "tr" denotes a matrix trace, as opposed to a thermal trace \( \text{Tr} \).

A major drawback, however, of the activity expansion for computing the canonical trace is that for \( N_\nu \approx N_\nu/2 \) (i.e., near half filling), \( \zeta_{N_\nu}(\sigma) \) involves a sum of terms that are large in magnitude and have alternating signs. In practical terms, the activity expansion is unstable in the mid-shell region because these terms cancel in the sum to 10–14 orders of magnitude, and there is a loss of numerical precision in the evaluation of Eq. (18). In actual calculations, we have found the activity expansion to be stable only for \( N_\nu \leq 4 \), or, using an equivalent hole formalism, for \( N_\nu \geq N_\nu - 4 \).

An alternative procedure for computing the canonical trace is to use Fourier extraction [11]. Starting from the grand-canonical trace, and defining \( \phi_m = 2\pi m/N_\nu \), we may write

\[ \det \left[ 1 + e^{i\phi_m e^{\beta \mu} U_\sigma} \right] = \sum_{N=0}^{N_\nu} e^{i\phi_m N} e^{\beta\mu N} \zeta_N(\sigma), \]  

(52)

where \( \mu \) is a parameter introduced to insure numerical stability throughout the range of particles in the model space, and is given below. Using the identity

\[ \frac{1}{N_\nu} \sum_{m=1}^{N_\nu} e^{i\phi_m K} = \delta_{K0}, \]  

(53)
valid for integer \( K \), we find

\[
\zeta_{N_z}(\sigma) = \frac{1}{N_z} \sum_{m=1}^{N_z} e^{-i\phi_m N_z} e^{-\beta \mu N_z} \det \left[ 1 + e^{i\phi_m} e^{\beta \mu U_\sigma} \right].
\]  

(54)

The expectation values of the one- and two-body density operators can be computed in a similar fashion:

\[
\langle a^\dagger_{a} a_{\beta} \rangle_{\sigma, N_z} = \frac{1}{N_z \zeta_{N_z}(\sigma)} \sum_{m=1}^{N_z} e^{-iN_z \phi_m} e^{-\beta N_z \mu \eta_m(\sigma)} \gamma^m_{\alpha\beta}(\sigma),
\]

(55)

and

\[
\langle a^\dagger_{a} a_{\beta} a^\dagger_{\delta} a_{\delta} \rangle_{\sigma, N_z} = \frac{1}{N_z \zeta_{N_z}(\sigma)} \sum_{m=1}^{N_z} e^{-iN_z \phi_m} e^{-\beta N_z \mu \eta_m(\sigma)} \times \left[ \gamma^m_{\alpha\beta}(\sigma)\gamma^m_{\gamma\delta}(\sigma) - \gamma^m_{\alpha\delta}(\sigma)\gamma^m_{\gamma\beta}(\sigma) + \delta_{\beta\gamma} \gamma^m_{\alpha\delta}(\sigma) \right],
\]

(56)

where

\[
\eta_m(\sigma) = \det \left[ 1 + e^{i\phi_m} e^{\beta \mu U_\sigma} \right]
\]

(57)

and

\[
\gamma^m_{\alpha\delta}(\sigma) = \left[ (1 + e^{i\phi_m} e^{\beta \mu U_\sigma})^{-1} e^{i\phi_m} e^{\beta \mu U_\sigma} \right]_{\beta\alpha}.
\]

(58)

The observables in Eq. (19–21) are, of course, independent of the value of \( \mu \) chosen. However, as the \( \zeta_\alpha \) vary rapidly with \( N \), a good choice for \( \mu \) is one for which the sum in Eq. (17) peaks at \( N = N_0 \). In order to find a good choice for \( \mu \), we first find the \( N_z \) eigenvalues \( \lambda_i \) of \( U_\sigma \), where \( i = 1, \ldots, N_z \), and \( | \lambda_1 | < | \lambda_2 | < \cdots \) (note that each eigenvalue has the form \( | \lambda_i | = \exp(-\beta \epsilon_i) \)). Thus, for the valence particles, we define \( \mu \) by

\[
| \lambda_{N_z} \lambda_{N_z+1} |^{1/2} = \exp \left[ -\beta \frac{\text{Re}(\epsilon_{N_z}) + \text{Re}(\epsilon_{N_z+1})}{2} \right] = \exp(-\beta \mu).
\]

(59)

This prescription allows us to use Fourier extraction for all even-even nuclei in both the \( sd \) and \( fp \) shells.

At first glance, the Fourier method appears to add substantial computational effort since the computation of a determinant scales as \( N_z^3 \) and it must be computed \( N_z \) times in Eq. (54). In fact, the computation of \( \zeta_{N_z} \) can be simplified considerably by computing the \( N_z \) eigenvalues, \( \lambda_i \), of \( U_\sigma \), in terms of which, the factor \( \eta_m(\sigma) \) can be written as

\[
\eta_m(\sigma) = \prod_{i=1}^{N_z} (1 + e^{i\phi_m} e^{\beta \mu \lambda_i}).
\]

(60)
In addition, the matrix $\gamma^m_{ab}(\sigma)$ is given by
\begin{equation}
\gamma^m_{ab}(\sigma) = \sum_\epsilon P_{\beta\epsilon}(1 + e^{i\phi_m} e^{B\epsilon} e^{i\phi_m} P^{-1}_{\beta\epsilon}),
\end{equation}
where $P$ is the transformation matrix associated with the diagonalization of $U_\sigma$.

### 3.4 Sign of the Monte Carlo Weight Function

The most important aspect of evaluating an integral using Monte Carlo methods is the sign of the weight function, $W(\sigma)$, as denoted by $\Phi(\sigma_k)$ in Eq.(22). Indeed, the entire method hinges on the fact that $W(\sigma)$ may be interpreted as a probability density, and, hence, must be positive definite everywhere. Unfortunately, in the Monte Carlo techniques discussed in these lectures, the weight function $W(\sigma) = G(\sigma)\zeta(\sigma)$ is not positive definite in the most general case. Indeed, as it turns out, for realistic interactions the average value for the sign $\langle \Phi \rangle$ used to normalize Eq.(22) is usually consistent with zero, which leads to large uncontrollable uncertainties in the Monte Carlo procedure.

Given the importance of the sign in the method, it is prudent to discuss the conditions that define the sign of the trace $\zeta(\sigma)$.

We begin by defining the ‘time-reversed’ partner of each creation and annihilation operator as
\begin{align}
\tilde{a}_{jm} &\equiv a_{jm} = (-1)^{j+m} a_{j-m} \\
\tilde{a}_{jm}^\dagger &\equiv a_{jm}^\dagger = (-1)^{j+m} a_{j-m}^\dagger.
\end{align}

Note that because of half-integer spins, we have
\begin{equation}
\tilde{a}_{jm} = -a_{jm}.
\end{equation}

As we will shortly see, the class of interactions that have positive-definite weight functions are of the form
\begin{equation}
\hat{H} = \sum_\alpha \left( \epsilon_\alpha \rho_\alpha + \epsilon_\alpha^* \bar{\rho}_\alpha - \frac{1}{2} \sum_\alpha \chi_\alpha \rho_\alpha \bar{\rho}_\alpha \right),
\end{equation}
where $\chi_\alpha > 0$, $\epsilon_\alpha$ can, in general, be complex, and $\rho_\alpha$ is a general density operator of the form
\begin{equation}
\rho_\alpha = \sum_{\gamma_\delta} C_{\gamma_\delta} a_{\gamma_\delta}^\dagger a_{\delta}.
\end{equation}
Note that in terms of $K$-coupled density operators, we have

$$\tilde{\rho}_a^{KM} = (-1)^K \rho_a^{K-M}. \tag{67}$$

In addition, the requirement of a negative coupling in Eq.(65) leads to a sign rule for the eigenvalues of the $E_K$'s, namely

$$\text{Sign}(V_a^K) = (-1)^{K+1}. \tag{68}$$

After applying the Hubbard-Stratonovich transformation to a Hamiltonian of the form Eq.(65), the effective one-body Hamiltonian becomes

$$\hat{h}(\sigma, \tau) = \sum_a (\epsilon_\alpha \rho_\alpha + \epsilon_\alpha^* \tilde{\rho}_\alpha - \chi_\alpha [(\sigma_\alpha + i \tau_\alpha) \rho_\alpha + (\sigma_\alpha - i \tau_\alpha) \tilde{\rho}_\alpha]), \tag{69}$$

so that $\rho_\alpha$ and $\tilde{\rho}_\alpha$ couple to complex-conjugate fields. We now wish to examine the form of the Hamiltonian matrix $h(\sigma \tau)$. We begin by ordering the single-particle state labels $n \ell jm$ so that those states with $m > 0$ are contained in the first half and their time-reversed partner in the second half. Now, in the case where all the $\chi_\alpha > 0$, the Hamiltonian matrix at the $i^{th}$ time slice will have the form

$$h_i = \begin{pmatrix} A_i & B_i \\ -B_i^* & A_i^* \end{pmatrix}, \tag{70}$$

and it is easy to verify that the imaginary-time propagator will be of the form

$$U = \prod_i \exp(-\Delta \beta h_i) = \begin{pmatrix} P & Q \\ -Q^* & P^* \end{pmatrix}. \tag{71}$$

Because of the structure of Eq.(71), the eigenvalues of $U$ come in complex-conjugate pairs, $\lambda$ and $\lambda^*$, with respective eigenvectors $|u\rangle$ and $|v\rangle$.

For the grand-canonical ensemble, the overlap is given by

$$\zeta = \det \left[ 1 + \begin{pmatrix} P & Q \\ -Q^* & P^* \end{pmatrix} \right] = \prod_i (1 + \lambda_i)(1 + \lambda_i^*) > 0. \tag{72}$$

If only particle-type (proton-neutron) conserving operators are present in Eq.(65), the overlap separates into two distinct terms for each nucleon type, and the total overlap is positive-definite as well.

In the zero-temperature formalism, if the trial wave function is comprised of an even number of particles and is chosen to consist of time-reversed pairs of single-particle states, so that we may write

$$\Psi_t = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{73}$$
where \( a \) and \( b \) are matrices with dimension \( N_e/2 \times N_e/2 \), then \( \Psi^\dagger U \Psi \) is a \( N_e \times N_e \) matrix of the form Eq.(71). The overlap function then satisfies

\[
\zeta = \det \left[ \Psi^\dagger U \Psi \right] > 0. \tag{74}
\]

In the canonical ensemble for \( N \) particles, a fixed number trace is involved, and, therefore,

\[
\zeta = \text{Tr}_N \left( \prod_i \exp(-\Delta \chi_i) \right) = \sum_{i_1 \neq i_2 \neq \ldots \neq i_N} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_N}. \tag{75}
\]

Although no rigorous proof exists, it has been found empirically that \( \zeta > 0 \) for even-even systems. In addition, for a Hamiltonian comprised of isoscalar density operators, i.e. \((\rho^p + \rho^n)\), \( \zeta \) has also been found to be positive definite for odd-odd \( N = Z \) systems. The reason for this is that the proton and neutron density operators in \( \hat{h} \) couple to the same fields, and, therefore, the proton and neutron traces are identical.

Often the third-component of the angular momentum is fixed by adding a Lagrange multiplier term \( \omega \hat{J}_z \) to the Hamiltonian. It should be noted, however, that the operator \( \hat{J}_z \) violates the time-reversal symmetry in \( \hat{h} \) required for the overlap to be positive definite, as \( \hat{J}_z = -\hat{J}_z \). On the other hand, it should be noticed by that cranking with an imaginary Lagrange multiplier \( i\omega \hat{J}_z \), the time-reversal properties are satisfied and the Monte Carlo sign is always positive definite.

In summary, the sign of the Monte Carlo weight function is positive definite for Hamiltonians of the form Eq.(65), where the couplings \( \chi_n \) are all positive. As it turns out, an important class of nuclear Hamiltonians do, indeed, satisfy this condition. Namely, those comprised of pairing plus separable multipole-multipole two-body interactions (such as quadrupole-quadrupole). Generally speaking, any “realistic” shell-model Hamiltonian is found to have a large component comprised of interactions of this type. In some sense, this encourages us to examine Hamiltonians possessing no sign problems, with the hope that it might be possible to extrapolate to fully realistic shell-model Hamiltonians. In Section 5, we will see the Monte Carlo approach applied to a separable interaction for \( sd \)-shell nuclei, and compared with exact results to demonstrate the feasibility of the method. In Section 6, we will see a method that will extrapolate to the “full” Hamiltonian, giving us the expectation values of operators in the ground state.
3.5 Approximate Solutions

Given the large dimensions of the integral, it might be beneficial to also consider approximate solutions to Eq.(22). The first approximate solution is to compute the observables at the maximum of the Monte Carlo weight function. In the case where only the direct part of the two-body matrix elements are used in the density-density decomposition, this equivalent to finding the Hartree mean field. With antisymmetrized two-body matrix elements, the auxiliary fields associated with the maximum of \( W \) are actually one half the Hartree and Fock fields [38]. An improvement upon the mean-field solution is to expand the exponent of the Monte Carlo weight function about the maximum up to quadratic terms, giving rise to RPA corrections to the mean field. In order to account for large-amplitude fluctuations of the mean field, one may integrate Eq.(22) using just one time slice. This approximation is reasonable for high temperatures, and is referred to as the static-path approximation (SPA) and has been used in nuclear physics for the pairing plus quadrupole-quadrupole interaction [39]. A correction to the static-path approximation is to again expand the exponent of the weight function up to second order about the static path, i.e. RPA corrections to the static path (SPA+RPA) [40]. It is important to note, however, that each of these approximations fail below some critical temperature, and to achieve the exact ground-state results, it is necessary to perform the multi-dimensional integral Eq.(22).

3.6 Monte Carlo Sampling

As previously noted, the dimensions of the integral Eq.(22) are too large to be evaluated using any form of quadrature rule, and must be evaluated using Monte Carlo methods. Central to a Monte Carlo evaluation of a multi-dimensional integral is the method for sampling the \( \sigma^i \) of Eq.(22), as they must be statistically independent. The sampling method used thus far for the Monte Carlo shell model is the method of Metropolis et al. [41]. In this method, we start with some set of fields \( \sigma^i_\alpha \), and choose a new \( i^{th} + 1 \) set by

\[
\sigma^{i+1}_\alpha = \sigma^i_\alpha + R(-1, 1)\Delta \sigma_\alpha,
\]

(76)

where \( R(X, Y) \) is a random number between \( X \) and \( Y \), and \( \Delta \sigma_\alpha \) is a pre-chosen increment. The new set of fields \( \sigma^{i+1}_\alpha \) are accepted if the ratio of weight functions satisfies

\[
W(\sigma^{i+1})/W(\sigma^i) \geq R(0, 1).
\]

(77)
The increment $\Delta \sigma_i$ is then chosen so that the ratio of accepted to rejected trials is approximately 50%.

The procedure used to select the $i^{th}$ + 1 fields in the Monte Carlo integration is to "sweep" through each of the time slices, and update the fields separately. Thus, one begins with the $i^{th}$ fields, and propagates from left-to-right in Eq.(25), i.e. from $N_t \rightarrow 1$. The fields are then updated for the first time slice, and the weight function is computed leaving the fields on all the other time slices unchanged. The updated fields on the first time slice are then accepted or rejected according to Eq.(77). The sweep then proceeds successively through all the time slices, propagating from right-to-left, and repeating the acceptance/rejection procedure at each time step.

One important point to be noted is that the observable sum must be computed with statistically independent samples. Clearly, successive Metropolis points are correlated since the $i^{th}$ + 1 point is dependent on the $i^{th}$ point. As such, the Monte Carlo observable should be evaluated using Metropolis points that are separated by $N_{\text{corr}}$ sweeps, where $N_{\text{corr}}$ is the Metropolis correlation length. Unfortunately, $N_{\text{corr}}$ is a quantity that grows with the dimensions of the integral. This is because as the dimensions increase, the increment $\Delta \sigma_i$ must become smaller so that the acceptance ratio will be approximately 50%. In early calculations, it was found that for the $sd$ shell the correlation length was of the order 100-300, whereas, in the $fp$ shell, the correlation length was $\approx$ 500. For calculations in the rare-earth region, with dimensions of the order 100000, the correlation length was estimated to be several thousand.

The simplest solution to the very long correlation length found with Metropolis sampling of the continuous integral is to recognize that the Hubbard-Stratonovich transformation as implemented in Eq.(13) is accurate only up to $(\Delta \beta)$ terms. Therefore, instead of performing a continuous integral, it might be more efficient to approximate Eq.(12) with a finite sum, namely

$$e^{\frac{1}{2} \Delta \sigma^2} = \sum_i w_i e^{\frac{1}{2} \frac{1}{2} |\Delta \sigma_i|^2 + \omega_i \Delta \sigma_i}.$$  

Indeed, a three-point Gauss-Hermite quadrature rule properly integrates Eq.(12) up to the third moment. The Metropolis algorithm for choosing the $i^{th}$ + 1 point is then modified in the following way. Instead of incrementing each sigma field by $\Delta \sigma$, we now choose a relative probability that any given field will change, $P(\sigma)$. We then choose a random number $R(0, 1)$ for each field, and if $R(0, 1) \leq P(\sigma)$, the sigma field is allowed to assume any of the permissible values with relative weight $w_i$. The advantage of a discrete
sum for the Hubbard-Stratonovich transformation is that with the modified Metropolis sampling described, the successive sweeps decorrelate much more quickly. Indeed, it is found that the correlation length with a three-point formula is of the order 10-20. The caveat, however, is that with the three-point formula, more time slices might be required in order achieve convergence in $\Delta \beta$. Having to use more time slices, however, is more than compensated by the shorter correlation length. The three-point Gauss-Hermite formula for Eq.(12) is being used in all current applications of the Monte Carlo method [10, 12].

4 Computational Considerations

In recent years, the concept of parallel computing has become a powerful tool in theoretical physics. To a large degree, the Monte Carlo approach described here is perfectly suited for parallel computation. This is because the Monte Carlo sum of Eq.(22) requires statistically independent samples. This can easily be achieved by “farming” the calculation out to several nodes, and thereby getting a substantial computational enhancement. This is especially true since the entire calculation can reside on each of the individual nodes, and very little data transfer is required. Indeed, the only time data transfer occurs is at the end of each sampling sweep, where the results are collected off of each node.

The Monte Carlo shell model code has been ported to several parallel platforms, among them: the 512-node Delta and Paragon at Caltech, the 512-node Paragon at Oak Ridge National Laboratory, the 28-node supercomputer at RIKEN, and the 400 node SP-2 at Phillips Laboratory, Air Force Material Command.

The computational work associated with the Monte Carlo shell model basically scales as the number of fields, and, therefore, as $(N^p_{f})^2 + (N^n_{e})^2$. Although this scaling is seemingly fairly rapid with $N_{e}$, it is far more gentle than the combinatorial growth of the number of basis states in the matrix-diagonalization approach, as is illustrated by Eq.(5). On an i860 based computer, such as the Delta and Paragon, the computational time required to complete an sd-shell calculation with approximately 3000 samples and 64 time slices is of the order 60-node hours, whereas a comparable calculation for fp-shell nuclei is of the order 300-node hours. One very important difference between the Monte Carlo and matrix-diagonalization approaches to the shell model is that with the Monte Carlo method, the computational time required for any given nucleus is more dependent on the number of single-particle states $N_{e}$ than on the number of valence particles $N_{v}$. As such, a calculation for $^{60}$Zn ($N_{v} = 10$) is not any more difficult than for $^{48}$Cr ($N_{v} = 4$). By
extrapolating the node-hours required for sd-shell and fp-shell nuclei, it is clear that full-space calculations in much larger model spaces are also feasible. Whereas, the traditional approach is currently limited to a full-space description of $^{50}$Mn ($N_e = 5$) in the fp shell.

One cautionary computational note with the AFMC method is that the propagator matrix $U$ is the product of the exponential of matrices whose diagonal elements may vary considerably. Indeed, the diagonal elements are primarily determined by the single-particle energies. For many Hamiltonians, the difference between the largest and smallest single-particle energies, $\Delta \epsilon$, is of the order 5-8 MeV. Hence the diagonal elements of $U$ may differ in scale by a factor of $e^{\beta \Delta \epsilon} \approx 3 \times 10^8$ for $\beta = 3$ MeV$^{-1}$. Such large differences in scale can lead to a loss in precision during the multiplication of the propagator matrices. Perhaps the only solution to this problem, which can be quite serious if large $\beta$ values are required, is the method of singular-value decomposition (SVD) [18, 42]. This algorithm is designed to preserve the scale in matrix-matrix multiplications so that high-precision can be maintained. An unfortunate feature, however, is that it increases the computational time considerably, since the SVD should be applied to the multiplication at each time slice. Efforts are underway to implement SVD in the current version of the Monte Carlo shell model program, and it appears that for static observables it might be possible to apply the SVD method every eighth time slice, allowing calculations to performed out to $\beta \sim 6$ MeV$^{-1}$ [43].

5 Demonstration with Separable Interactions

In this section, results obtained with the auxiliary-field Monte Carlo shell model (AFM-CSM) using a schematic interaction that has a sign of unity are presented and compared with those obtained from the direct diagonalization method. The purpose is, of course, to demonstrate the feasibility of the method, so that we may begin to have confidence in the algorithm to tackle problems not possible within the traditional approach.

The interaction was chosen to be of the form pairing plus multipole given by citeafmc5

$$\hat{H} = \sum_i \epsilon_i \sqrt{2j_i + \frac{1}{2}} \rho_{ii}^{00,i=0} - G \sum_{ij} [A_{j=0,T=1}^j (ii) \otimes A_{j=0,T=1}^j (jj)]^{J=0,T=0}$$

$$- \frac{1}{2} \sum_{K=0,2,4} \chi_K (-1)^M \rho^{K,M,t=0} \rho^{K-M,t=0},$$

where $G > 0$ and each $\chi_K > 0$. The parameters were chosen to reproduce parts of the sd-shell Hamiltonian of Wildenthal [2], and can be obtained from the author upon request.
For each of the calculations presented in this section, the value \( \Delta \beta = 0.0625 \text{ MeV}^{-1} \) was used and approximately 2000 Monte Carlo samples were collected. Being amongst the first calculations, they were performed using the continuous form of Eq. (12), and the independence of the individual samples was tested by computing the auto-correlation function for \( \langle H \rangle_c \). As such, \( N_{\text{corr}} \) was found to be \( \approx 150 \).

5.1 Static Observables

We begin with a compendium of results for the Ne isotopes. Shown in Fig. 2 are the AFMCSM expectation values (open symbols) (a) \( \langle \hat{H} \rangle \), (b) \( \langle Q^2 \rangle \) (quadrupole moment), (c) \( \langle J^2 \rangle \) (angular momentum), and (d) \( \langle T^2 \rangle \) (isospin) as functions of \( \beta \) for even-even Ne isotopes. Exact calculations within the canonical ensemble for \(^{20,22}\text{Ne}\) using the eigenvalues obtained from the shell-model code OXBASH \(^{28}\) are indicated by the curves in the figure. In addition, the ground-state observables are plotted using solid symbols at \( \beta = 2.5 \text{ MeV}^{-1} \). Generally, we see that the Monte Carlo procedure is in good agreement with the exact calculations. Shown in Fig. 3, are the results for the same quantities for the odd-odd \( N = Z \) nucleus \(^{23}\text{Na}\).

We may also study rotating nuclei using the cranking Hamiltonian, \( \hat{H}' = \hat{H} - \omega \hat{J}_z \). The systematics for cranking \(^{22}\text{Ne}\) are shown in Fig. 4, where we display \( \langle \hat{H} \rangle \), \( \langle \hat{J}_z \rangle \), and the sign \( \langle \Phi \rangle \) as a function of \( \beta \) and \( \omega \). We find that the sign decays rapidly as both the cranking frequency and \( \beta \) increase. The maximum \( J_z \) available to \(^{22}\text{Ne}\) is 10, and, therefore, the \( \omega = 2 \) case can be considered as an extreme limit.

5.2 Nuclear Shapes

It is of particular interest to determine the quadrupole shape of a nucleus as function of temperature and angular momentum. It is generally expected that some nuclei may exhibit a sudden phase transition from a prolate to spherical shape as the temperature increases \(^{44}\). In addition, within the framework of the cranking model, deformed nuclei are expected to exhibit a transition from prolate to oblate ellipsoids as the cranking frequency (or angular momentum) increases \(^{45}\).

One measure of the quadrupole deformation is the expectation value of \( Q^2 \). As is illustrated in the proceeding section, \( \langle Q^2 \rangle \) is considerably larger for nuclei that are expected to exhibit prolate deformations such as \(^{20,22,24}\text{Ne}\), and is much smaller for spherical nuclei such as \(^{26}\text{Ne}\). The interpretation of shapes from \( \langle Q^2 \rangle \), however, suffers from two short-
Figure 2: The results of the Monte Carlo calculation for the expectation values of (a) $\langle \hat{H} \rangle$, (b) $\langle q^2 \rangle$, (c) $\langle J^2 \rangle$, and (d) $\langle T^2 \rangle$ as a function of $\beta$ for $^{20}\text{Ne}$ (circles), $^{22}\text{Ne}$ (squares), $^{24}\text{Ne}$ (diamonds), and $^{26}\text{Ne}$ (triangles). Where absent, the error bars are smaller than the size of the symbols. The solid ($^{20}\text{Ne}$) and dashed-dot ($^{22}\text{Ne}$) lines indicate the canonical results obtained from the eigenvalues of an exact diagonalization. The ground-state expectation value for each nucleus is plotted with a solid symbol at $\beta = 2.5 \text{ MeV}^{-1}$ (except for $J^2$, for which the ground-state value is zero for all nuclei).

comings. First, $Q^2$ contains a one-body term proportional to $\langle r^4 \rangle$, which is present even for spherical nuclei, and tends to obscure the contribution due to the nuclear deformation. In addition, $\langle Q^2 \rangle$ does not distinguish between prolate and oblate shapes.

In order to obtain a more detailed picture of the deformation, we examine the components of the quadrupole operator $Q_\mu = r^2 Y_{2\mu}^*$. Note, however, that due to rotational invariance of the Hamiltonian, the expectation value of any component $Q_\mu$ is expected to vanish. On the other hand, for each Monte Carlo sample, there exists an intrinsic frame in which it is possible to compute the three non-zero components $Q_0^\prime$, $Q_2^\prime$, and $Q_{-2}^\prime$ (the prime is used to denote the intrinsic frame). The intrinsic quadrupole moments can then
be related to the standard deformation coordinates $\beta$ and $\gamma$ [46] by

\[
\langle Q'_0 \rangle_\sigma = \frac{3}{2\pi} \sqrt{\frac{4\pi}{5}} \langle r^2 \rangle_\sigma \beta_\sigma \cos \gamma_\sigma \\
\langle Q'_2 \rangle_\sigma = \frac{3}{2\pi} \sqrt{\frac{4\pi}{5}} \langle r^2 \rangle_\sigma \frac{\beta_\sigma}{\sqrt{2}} \sin \gamma_\sigma \\
\langle Q'_{-2} \rangle_\sigma = \frac{3}{2\pi} \sqrt{\frac{4\pi}{5}} \langle r^2 \rangle_\sigma \frac{\beta_\sigma}{\sqrt{2}} \sin \gamma_\sigma.
\]

The task remains to compute the quadrupole moments in the intrinsic frame for each Monte Carlo sample. This is accomplished by computing and diagonalizing the expectation value of the cartesian quadrupole tensor $Q_{ij} = 3x_i x_j - \delta_{ij} r^2$ for each Monte Carlo sample. From the three eigenvalues, it is straightforward to determine the deformation parameters as [47]

\[
\langle Q'_{11} \rangle_\sigma = \frac{2\pi}{5} \left( \sqrt{3} (\langle Q'_2 \rangle_\sigma + \langle Q'_{-2} \rangle_\sigma) - \sqrt{2} \langle Q'_0 \rangle_\sigma \right)
\]
Figure 4: Cranked results for $^{22}$Ne are given as a function of the cranking frequency $\omega$. Calculations were performed at $\beta = 0.5$ (circles), 1.0 (squares), 1.5 (diamonds), and 2.0 (triangles).

$$
(Q'_{22})_o = \sqrt[5]{\frac{2\pi}{3}} \left(-\sqrt{3}(Q'_{2})_o + (Q'_{-2})_o\right) - \sqrt{2}(Q'_{0})_o
$$
\[ \langle Q'_{33} \rangle_{\sigma} = 2\sqrt{\frac{4\pi}{5}} \langle Q_0' \rangle_{\sigma}. \] (81)

Note, from Eq. (80) one finds \( Q'_{22} \leq Q'_{11} \leq Q'_{33}. \)

To illustrate the intrinsic deformation, we plot in Fig. 5 the distribution function \( F(\beta, \gamma) = f(\beta, \gamma)\beta^4 \sin(3\gamma) \) for temperatures \( T = 0.5, 1.0, 2.0, \) and \( 4.0 \) MeV. The distribution function was computed from the set of the \( \beta \) and \( \gamma \) values from each Monte Carlo sample, and then smoothed with a symmetric Gaussian with a width of 0.01. Although the volume element \( \beta^4 \sin(3\gamma) \) tends to push the function towards \( \gamma = \pi/6 \), there is a definite trend from a prolate deformation at low temperature to a symmetric spherical shape at higher temperatures.

\[ \text{Figure 5: Distribution functions } F(\beta, \gamma) \text{ are shown for } ^{22}\text{Ne at different temperatures } T. \]
6 Realistic Interactions

In the previous section, the AFMCSM was demonstrated to be successful for the case of semi-realistic schematic interactions that possess the feature that the Monte Carlo weight function is always positive definite. Unfortunately, this is not always the case for the general Hamiltonian. Indeed, for most realistic Hamiltonians, the Monte Carlo sign degrades with the application of a few time slices, and quickly goes to zero. In this section, I will describe a method that can in principle give the correct results for Hamiltonians applicable in nuclear physics.

We begin by analyzing the behaviour of a realistic interaction in terms of the sign rule Eq.(68) as is illustrated in Fig. 6. Here the eigenvalues \( \Lambda^K_{\alpha} = (-1)^{K+1} V^K_{\alpha} \) for the fp-shell Hamiltonian of Richter et al. [48] are displayed. The first thing to be noted is that most of the large eigenvalues satisfy the sign rule, and, therefore, have a "good" Monte Carlo sign. As such, we can decompose the nuclear Hamiltonian into two parts

\[
\hat{H} = \hat{H}_G + \hat{H}_B,
\]

where \( \hat{H}_G \) satisfies the sign rule, and has "good" sign, while \( \hat{H}_B \) violates the sign rule, leading to negative signs in the Monte Carlo weight function, and, therefore, has a "bad" sign. It is then possible to construct another Hamiltonian, denoted \( \hat{H}_g \) by

\[
\hat{H}_g = \hat{H}_G + g \hat{H}_B.
\]

Note that if \( g < 0 \), then \( \hat{H}_g \) satisfies the Monte Carlo sign rule, and a Monte Carlo calculation is possible. The hope then, is that we may be able to compute the expectation value of various observables for the Hamiltonian \( \hat{H}_g \) and then extrapolate to the Hamiltonian of interest at \( g = 1 \).

As it turns out, the essential feature of \( \hat{H}_g \), with \( g < 0 \) is to enhance the pairing part of the interaction, as can be seen in Fig. 7, where the \( T = 1 \) matrix elements of the Hamiltonian are plotted for all \( J \) values and for both \( g = 0 \) and \( g = 1 \) (the true Hamiltonian). Note that the most significant change occurs in the pairing \( (J = 0) \) channel. It should be noted that the main feature of the density-density decomposition of the pairing interaction is that it is present in all possible \( K \)-values. As such, it is possible to transform any Hamiltonian into a Hamiltonian with a "good" sign by steadily increasing the pairing strength until all the eigenvalues obey Eq.(68).

One very important feature of the Hamiltonian \( \hat{H}_g \) is that at temperature \( T = 0 \) MeV
Figure 6: The eigenvalues $V_{a}$ of the Richter fp-shell interaction multiplied by the factor $(-1)^{K+1}$. The eigenvalues for each $K$ value are plotted in increasing order.

(or $\beta \to \infty$) $\hat{H}_{g}$ satisfies a variational principle. Namely

$$\langle \hat{H} \rangle_{g} \geq \langle \hat{H} \rangle_{g=1}, \quad (84)$$

where the subscript $g$ denotes a trace over the eigenstates of the $\hat{H}_{g}$. In addition, because of the variational principle, we also have

$$\frac{d \langle \hat{H} \rangle_{g}}{dg} \bigg|_{g=1} = 0. \quad (85)$$

Therefore, we may hope to obtain information regarding the true solution by computing the expectation value of observables for $g < 0$, and extrapolating with a smooth function to $g = 1$. It should be pointed out that Eqs. (84) and (85) are in no way a guarantee of success. The variational characteristic for the Hamiltonian only tells us that the true value of the expectation value of the Hamiltonian is less than the values obtained at each $g$ value. It is not a guarantee that $\langle \hat{H} \rangle_{g}$ will be a smooth function of $g$ or that it will be monotonic. Nonetheless, comparisons with exact results indicate that for the nuclear problem the proposed extrapolation method can be successful. Shown in the bottom panel of Fig. 8 is $\langle \hat{H} \rangle_{g}$ for an fp-shell calculation of $^{48}$Cr using the KB3 [34] interaction as a function of $g$. The exact calculation [35] is represented by the open circles and the dotted line. The Monte Carlo results are shown for $\beta = 2$ MeV$^{-1}$ and $\Delta \beta = 1/32$ MeV$^{-1}$, and are indicated by the solid symbols (with error bars) and connected with a solid line. The point at $g = 1$ is obtained by extrapolating the Monte Carlo results with a polynomial satisfying Eq. (85). Note that the deviation between the Monte Carlo and exact results is
Figure 7: The two-body matrix elements $V_{T=1}(ij, kl)$ of the Richter fp-shell interaction (solid circles) and its "good" part (open circles), for $J \leq 4$ The ordering for each $J$ is arbitrary.

different as a function of $g$, indicating that an extrapolation on $\Delta \beta$ for each $g$ value is warranted before extrapolating to $g = 1$.

In the upper two panels of Fig. 8, are comparisons between the exact and Monte Carlo results as a function of $g$ for the total Gamow-Teller strength and the total $E(2)$ strength. Note that for these operators there is no variational principle, and for both operators there is significant curvature in the region $g > 0$, where a Monte Carlo calculation is
Figure 8: Comparison between Monte Carlo (solid circles with error bars) and exact diagonalization (open circles) results for the Hamiltonian $\hat{H}$, the total Gamow-Teller strength $B(GT^\pm)$, and the total $BE(2)$ strength as a function of $g$ for $^{48}\text{Cr}$ using the KB3 interaction. The solid line represents a polynomial extrapolation of the various quantities.

...not possible. Nonetheless, the Monte Carlo shell model with the extrapolation procedure...
reproduces the exact result to within the uncertainty of the extrapolation.

To conclude this section, results obtained with the Monte Carlo shell model for the nucleus $^{54}$Fe [10] using the full fp-shell and the Hamiltonian of the Richter et al. [48] are shown in Fig. 9. The solid lines in the figure represent polynomial extrapolations of the various quantities, and the extrapolated results are plotted at $g = 1$ along with the associated uncertainty. Note that the number of basis states for this nucleus as estimated with Eq.(5) is of the order $4.9 \times 10^9$, and it is not possible to perform a full-space shell-model calculation for this nucleus with any shell-model code currently available. The calculations were performed on the 512-node Delta and Paragon at Caltech with $\beta = 2.0$ MeV$^{-1}$ and $\Delta \beta = 1/16$ MeV$^{-1}$, and approximately 4000 samples were accumulated for each $g$ value.

7 Response Functions

In addition to static observables, the Monte Carlo approach can also be used to compute response functions for transition operators. We begin with the thermal trace

$$f_\beta(E) = Z^{-1} \sum_{i,j} e^{-\beta E_i} |\langle f| \hat{O}(i) \rangle|^2 \delta(E - (E_j - E_i))$$

$$= Z^{-1} \int_{-\infty}^{\infty} \frac{dt}{2\pi} e^{i E t} \text{Tr} \left[ e^{-\beta \hat{H}} \hat{O}(t) \hat{O}(0) \right].$$

Unfortunately, the presence of the real-time propagators in Eq.(87) leads to an uncontrolled Monte Carlo sign since the Hubbard-Stratonovich transformation must be applied to the time-evolution operator $e^{i \hat{H} t}$. It is thus advantageous to switch to an imaginary time $\tau = it$, so that the response function becomes the inverse Laplace transform of the imaginary-time auto-correlation function of the transition operator, namely

$$f_\beta(E) = Z^{-1} \int_{-\infty}^{\infty} \frac{d\tau}{2\pi i} e^{-E \tau} \text{Tr} \left[ e^{-\beta \hat{H}} \hat{O}(\tau) \hat{O}(0) \right] = \int_{-\infty}^{\infty} \frac{d\tau}{2\pi i} e^{-E \tau} R(\tau),$$

where $\hat{O}(\tau) = e^{\beta \tau} \hat{O} e^{-\beta \tau}$. The utility of Eq.(88) is that it can be evaluated using Eq.(17) and $\hat{O}(\tau) \hat{O}(0)$ can be discretized with the same time step as the operator $e^{-\beta \hat{H}}$.

Before proceeding with a discussion on performing the inverse Laplace transform to $R(\tau)$, it is useful to discuss some of its properties. First, the $n^{th}$ intrinsic moment of $f_\beta$ is given by

$$\langle f_\beta \rangle_n = \frac{d^n}{d\tau^n} \ln R(\tau) \bigg|_{\tau = 0}$$
Figure 9: The results of the Monte carlo calculations for $^{54}$Fe at $\beta = 2$ MeV$^{-1}$ for several observables as a function of $g \leq 0$. $Q = Q_p + Q_n$ is the isoscalar quadrupole, $Q_v = Q_p - Q_n$ is the isovector quadrupole, GT$^+$ is the Gamow-Teller operator changing a proton into a neutron, and $M1$ is the magnetic moment operator using free nucleon $g$ factors. The lines represent polynomial extrapolations; the extrapolated values and corresponding uncertainties are shown at $g = 1$.

(note $\langle E^n \rangle = R(0)^{-1} d^n R(\tau)/d\tau^n |_{\tau=0}$). Also, because of the thermal trace, $R(\tau)$ is sym-
metric about $\tau = \beta/2$ for hermitian transition operators such as $E_1$, $M_1$, and $E_2$. Indeed, in the limit that $f_\beta$ involves only one transition, $R(\tau) \propto \cosh[(E_f - E_i)(\tau - \beta/2)]$. As a consequence of the symmetry about $\beta/2$, the smallest energy that can be resolved in $f_\beta$ is of the order $2/\beta$. Further, since $R(\tau)$ decays exponentially for short times, the largest energy that can be resolved for $f_\beta$ is the inverse of the discretization size, namely $1/\Delta \beta$.

Generally speaking, the inverse Laplace transform in Eq.(88) is an ill-posed problem, and is usually performed using Maximum Entropy (MaxEnt) reconstruction techniques. The details and the philosophy behind MaxEnt reconstruction lie beyond the scope of these lectures, and will, therefore, only be briefly outlined here. The reader interested in further details is referred to Refs. [49, 50]. We begin, for illustrative purposes, with a simplified technique that works well when most of the strength is concentrated into a single peak. We assume that $f_\beta$ consists of a single line whose shape is of the form

$$f_\beta(E) = \frac{f_0 e^{\beta/2(E-E_0)}}{\Delta E \sqrt{2\pi}} \left[ e^{-(E-E_0)^2/2\Delta E^2} + e^{-(E+E_0)^2/2\Delta E^2} \right].$$

The Laplace transform of $f_\beta$ is then

$$R_f(\tau) = f_0 \left[ e^{-\tau E_0} + e^{(\tau-\beta)E_0} \right] \exp \left( \frac{1}{2}(\beta/2 - \tau)^2 \Delta E^2 \right).$$

In this limit, it is possible to perform a least-squares fit of $R_f$ to $R(\tau)$ to obtain the centroid $E_0$ and width $\Delta E$. The strength $f_0$ can then be determined by the total strength.

For a MaxEnt reconstruction, we begin by discretizing $f_\beta$ by using $N_t$ equally spaced points with maximum and minimum energy of $1/\Delta \beta$ and $2/\beta$, respectively ($\Delta E = (1/\Delta \beta - 2/\beta)/N_t$). Then, $f_\beta$ assumes the form

$$f_\beta(E) = \sum_i f_i e^{\beta/2(E-E_i)} \frac{1}{\Delta E \sqrt{2\pi}} \left[ e^{-(E-E_i)^2/2\Delta E^2} + e^{-(E+E_i)^2/2\Delta E^2} \right],$$

and

$$R_f(\tau) = \sum_i f_i \left[ e^{-\tau E_i} + e^{(\tau-\beta)E_i} \right] \exp \left( \frac{1}{2}(\beta/2 - \tau)^2 \Delta E^2 \right).$$

The goal of the MaxEnt method is to compare $R_f(\tau)$ with the computed $R(\tau)$ so as to find the most probable (or mean) values for the set $\{f_i\}$. For this, the probability for $\{f_i\}$ is taken to be

$$Pr(\{f_i\}) = (Z_5 Z_L)^{-1} \exp \left( \alpha S - \frac{1}{2} \chi^2 \right),$$

with the entropy $S$ defined as [49]

$$S = \sum_i [f_i - m_i - f_i \ln(f_i/m_i)],$$
where \( \{m_i\} \) is a default model (usually chosen to be uniform), \( \chi^2 \) is given by [49]

\[
\chi^2 = \sum_i [R_i(\tau_i) - R(\tau)] [R_i(\tau_j) - R(\tau)] / \sigma_i \sigma_j,
\]  

(96)

where \( \sigma_i \) is the Monte Carlo uncertainty associated with each \( R(\tau) \), and

\[
Z_S = \int \prod_i d\tau_i f_i^{-1/2} e^{\alpha S}
\]  

(97)

and

\[
Z_L = \int \prod dR(\tau_i) e^{-\chi^2/2}
\]  

(98)

are normalizations for the entropy and \( \chi^2 \) distributions. Note that for strength functions with considerable structure, the single-line least-squares fit outlined above can be used to determine an improved default model \( \{m_i\} \). Finally, the parameter \( \alpha \) is chosen so as to maximize the conditional probability of \( \alpha \) given the data set \( \{R(\tau_i)\} \) and the default model \( \{m_i\} \)

\[
Pr(\alpha|R(\tau_i)\{m_i\}) \propto (Z_L Z_S)^{-1} \int \prod_i d\tau_i f_i^{-1/2} \exp(\alpha S - \frac{1}{2} \chi^2).
\]  

(99)

For the most part, \( Pr(\alpha|R(\tau_i)\{m_i\}) \) is a sharply peaked function of \( \alpha \). As such, Eq.(97) can be evaluated using a saddle-point approximation (also for \( Z_S \)).

For demonstration purposes, the results for \(^{22}\text{Ne}\) using the schematic interaction of Eq.(79) in Section 5. at \( \beta = 2.0 \text{ MeV}^{-1} \) (16 and 32 time slices) are displayed in Fig. 10 for: a) the isoscalar quadrupole, \( \langle Q^0(\tau) \cdot Q^0(0) \rangle \); b) the isovector quadrupole, \( \langle Q^1(\tau) \cdot Q^1(0) \rangle \); and c) the isovector angular momentum, \( \langle J^1(\tau) \cdot J^1(0) \rangle \). Generally, the reconstructed strength functions are in good agreement with the exact results, especially in those cases where most of the strength is concentrated in a single peak as for the isoscalar quadrupole. For the situation in which the strength function is strongly fragmented, as in the case for the isovector angular momentum, the various lines can be reconstructed only by using many more time slices so that sufficient information in the small \( \tau \) region of the imaginary-time response function exists. It is clear that in this case it is necessary to disentangle several decaying exponentials with different slopes.

8 Finite Temperature

At times it is useful to examine the properties of nuclei at a finite excitation energy above the ground state. One difficulty, however, is that the level density in nuclei increases
Figure 10: Response functions for $^{22}$Ne are shown, along with MaxEnt reconstruction of the strength functions. The calculated response functions (left) of the isoscalar quadrupole (top), isovector quadrupole (middle), and isovector angular momentum (bottom) are shown for $\Delta \beta = 0.125$ (circles), and $\Delta \beta = 0.0625$ (triangles). MaxEnt reconstruction of the strength functions are also given for $\Delta \beta = 0.125$ (dotted line) and $\Delta \beta = 0.0625$ (dashed line). Exact results given as the delta function peaks.

dramatically with increasing excitation energy. As such, it may not be particularly useful
to examine the properties of individual levels. In addition, in large model spaces, it may be extremely difficult to isolate every eigenvalue. An alternative procedure is to perform a sum over all the states in the model space, and by the introduction of a Lagrange multiplier $\beta$ constrain the expectation value of the Hamiltonian. This is the basis of statistical mechanics, and is embodied in the partition function as described in Eq. (7) of Section 3. We immediately see that AFMC techniques are well suited to the calculation of thermal observables. In this section, I will outline some of the general features of the applications of Monte Carlo approach at finite temperature. There are several studies currently underway, and in these lectures, I will only outline some of the main features and difficulties.

The basic thermal quantities of interest are

\[ Z(\beta) = \text{Tr} \left( e^{-\beta \hat{H}} \right) \]  
\[ \rho(E) = \int_{-\infty}^{\infty} \frac{d\beta}{2\pi} e^{-\beta E} Z(\beta) \]  
\[ F = -\beta \ln Z = U - TS \]  
\[ S = \frac{\partial F}{\partial \beta} \]  
\[ U = \langle \hat{H} \rangle = -\partial \ln Z / \partial \beta \]  
\[ C_V = = \frac{\partial U}{\partial \beta} = \beta^2 (\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2) = \beta^2 \partial^2 \ln Z / \partial \beta^2, \]

where $\rho(E)$ is the level density as a function of excitation energy $E$, $F$ is the free energy, $S$ the entropy, $U$ the internal energy, and $C_V$ the specific heat. The two quantities of primary interest are the level density and the specific heat, which is interesting because sharp peaks in $C_V$ are an indication of the presence of first-order phase transitions. Phase transitions that might be observable in nuclei are a shape transition that should occur in strongly deformed nuclei at $T \approx 1.5$ MeV [44] and the dissappearance of the pairing gap [51].

As a demonstration of the usefulness of the AFMC method for finite temperature, shown in Fig. 11 are the results for an fp-shell calculation of $^{54}$Fe using a schematic interaction of the form Eq. (79). The calculations contain 4000 samples and were performed for $\Delta \beta = 1/32$ MeV$^{-1}$. In general, we see that the results for the Hamiltonian are smooth as a function of $\beta$. For the specific heat, however, there appears to be no evidence of a first-order phase transition, although there are fairly large fluctuations in $C_V$ for $0.5 \geq T \geq 1.0$. This is primarily caused by the fact that the calculations were performed in steps of constant $\Delta \beta$. As such, $C_V = -\beta^2 d\langle \hat{H} \rangle / d\beta$, and small fluctuations
in $d\langle \hat{H} \rangle / d\beta$ are amplified by the $\beta^2$ factor. For low temperatures, it is perhaps best to perform calculations with constant $\Delta T$. To some degree, these fluctuations may also be smoothed out by smoothing the partition function $Z(\beta)$, which we can obtain via

$$Z(\beta) = \int_0^\beta d\beta \langle \hat{H} \rangle + \ln Z(0),$$

(106)

where $Z(0)$ is the total number of Slater determinants in the space, and is given by Eq.(5). With the partition function thus computed, one can reconstruct the level density by performing the inverse Laplace transform in Eq.(99) using the maximum entropy techniques outlined in Section 7.

![Graphs showing Monte Carlo results as a function of $\beta$ and temperature $T = 1/\beta$ for $^{54}$Fe in the fp shell using a schematic interaction.](image)

Figure 11: Monte Carlo results as a function of $\beta$ and temperature $T = 1/\beta$ for $^{54}$Fe in the fp shell using a schematic interaction.

The situation for realistic interactions is somewhat more complicated as we must use the extrapolation procedure outlined in Section 6. An important consideration, however, is that the expectation value of $\hat{H}$ as a function of $g$ is no longer a variational quantity, and large errors in the extrapolation can occur. In addition, as was mentioned in Section 6, the primary feature of the $H_g$ Hamiltonian is a significant enhancement of the pairing part
of the interaction. Because of this, a rather large gap opens up between the $J = 0$ ground state and the excited states of the Hamiltonian $H_g$. For this reason, the extrapolated value of $\langle \hat{H} \rangle_g$ is generally suppressed relative to the exact result. In some sense, this is a good feature when computing ground-state observables, provided a level crossing does not occur, since it may not be necessary to carry the calculation out to very large values of $\beta$. In regards to finite temperature observables, however, this is a diasterous feature. Two methods to fix this problem are now currently under review. The first is to modify the good part in $H_g$, so that

$$H_g = [1 - (1 - g)/A]H_G + gH_B,$$  \hspace{1cm} (107)

where $A \approx 2$. In this form, the gap between the ground state and the excited states of the $H_g$ Hamiltonian is reduced considerably [43]. The second is a variational free energy method [52] based on

$$F_g(\beta) = \langle \hat{H} - \hat{H}_g \rangle_g - \ln Z_g(\beta),$$  \hspace{1cm} (108)

where

$$\ln Z_g(\beta) = \int_0^\beta d\beta \langle \hat{H}_g \rangle_g + \ln Z(0).$$  \hspace{1cm} (109)

Since $F_g$ is a variational quantity, we may now extrapolate to $g = 1$ using a function having zero derivative at $g = 1$.

9 Conclusions

In these lectures, I have outlined the main features of the auxiliary-field Monte Carlo approach to the nuclear shell model, otherwise known as AFMCSM. As such, I have included the main features of refs. [7, 8, 9, 10, 11, 12] into these lectures as a central reference on the method. We have seen that the AFMCSM can yield exact results in regions not possible with the traditional approach because of computational limitations. The primary difficulty in implementing the AFMCSM is the sign of the Monte Carlo weight function, which was found to be unity for an entire class of schematic interactions often used in nuclear physics. Unfortunately, realistic interactions possess a bad Monte Carlo sign, and, in effect, it is not possible to perform a direct calculation. Instead, we separate the true Hamiltonian into the components with good and bad sign, and multiply the bad components by the factor $g < 0$, so that a Monte Carlo calculation is feasible. Then, we extrapolate the $g < 0$ results to $g = 1$ to obtain the expectation value of observables for the true Hamiltonian. Also, because of the sign problem, the
AFMC method is found to work best for even-even and N=Z nuclei. One drawback of the AFMC procedure is that detailed spectroscopic information is not obtainable. However, by utilizing the strength function, information about excited states can be obtained. To conclude, the Monte Carlo approach to the shell model is a powerful technique that has the potential to add significantly to our understanding of nuclear structure since it permits us to perform full-space calculations with realistic interactions that were not possible until now.

10 Acknowledgements

I would like to thank Professor Hiro Sagawa for his kind hospitality and invitation to this workshop so that I could give these lectures. I would also like to thank W. Nazarevicz and M. Guidry for their critical reading of this manuscript. Theoretical nuclear physics research at the University of Tennessee is supported by the U.S. department of Energy through the contract No. DE–FG05–93ER40770. Oak Ridge National Laboratory is managed by Martin Marietta Energy Systems, Inc. for the U.S. Department of Energy under contract No. DE–AC05–84OR21400.

References


[43] D. J. Dean, private communication.


**DISCLAIMER**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.