LA-UR 90-3249

CONF-9008172--2

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

LA-UR--90-3249

DE91 000216

TITLE: APPLICATIONS OF MAXENT TO QUANTUM MONTE CARLO

AUTHOR(S): Richard N. Silver
Devinderjit S. Sivia
James E. Gubernatis
Mark Jarrell*

SUBMITTED TO: Maxent 90 Proceedings, to be published Proc. X
Maximum Entropy Workshop, Laramie 1990, EDS: W. T. Grandy, Jr;
L. H. Schick

*Department of Physics
The Ohio State University
Columbus, Ohio 43210

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2629 5/81

APPLICATIONS OF MAXENT TO QUANTUM MONTE CARLO

Richard N. Silver, Devinderjit S. Sivia, James E. Gubernatis
Theoretical Division and Manuel Lujan Jr. Neutron Scattering Center
MS B262 Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Mark Jarrell†
Department of Physics
The Ohio State University
Columbus, Ohio 43210

ABSTRACT: We consider the application of maximum entropy methods to the analysis of data produced by computer simulations. The focus is the calculation of the dynamical properties of quantum many-body systems by Monte Carlo methods, which is termed the "Analytic Continuation Problem." For the Anderson model of dilute magnetic impurities in metals, we obtain spectral functions and transport coefficients which obey "Kondo Universality."

Introduction

A new area for the application of maximum entropy and Bayesian methods is in computer simulations. We will provide an example of how data produced by stochastic methods for solving the many-particle Schrodinger equation may be analyzed by the maximum entropy method. Specifically, we will calculate the excitation spectra and transport coefficients of an important many-body system, the Anderson model (Anderson, 1961) for dilute magnetic impurities in alloys. This is an extremely ill-posed inverse problem similar to numerically inverting a Laplace transform from incomplete and noisy data. In our case, the data are computer generated, they are not statistically independent, and they are not necessarily Gaussian distributed. We will discuss how Classic MaxEnt (Gull, Skilling, 1989) noise scaling, error estimates on integrated quantities, and the choice of informative default model are all essential for obtaining good results.

Our work is the first example of what we expect will develop into a wide range of applications of maximum entropy and Bayesian methods to computer simulations of quantum many-body systems. These systems include normal and superfluid ^{4}He, nucleons in nuclei and nuclear matter in neutron stars, quantum chromodynamics in elementary particle physics, and strongly correlated electronic systems, e.g. high temperature superconductivity, magnetism, the quantum hall effect, heavy fermion materials, etc. What all these diverse systems have in common is that quantum effects are essential, the systems must be described by many-particle wave functions of the appropriate Bose or Fermi symmetry, and the interaction potentials between the particles are very strong. The strong potential implies that there is little hope for conventional theoretical approaches based on a perturbation expansion in powers of the potential starting from the noninteracting limit, e.g.

†Permanent address: Dept. of Physics, University of Cincinatti, ML 11, Cincinatti, OH 45221.

Feynman diagram expansions. The "small parameter" needed to justify such expansions does not exist for these systems. The essential physics often corresponds to an infinite power resummation of the perturbation expansion. Usually, it is feasible to calculate only a few terms in the perturbation expansion.

There are many alternatives to perturbation theory, each with its own domain of success. The most common approach is phenomenological; that is, one eschews a first principles (*ab initio*) understanding and, instead, one builds a model for the behavior with parameters determined from experiment. An example is the Fermi liquid theory for metals. An *ab initio* approach is to hypothesize approximate wave functions, which by a variational principle can be made to approach the ground state wave functions of the system. This has many successes for the static properties such as pair-correlation functions and momentum distributions, but it cannot describe the dynamical behavior such as collective modes and transport coefficients. The renormalization group approach focuses on properties, such as critical phenomena at phase transitions, which do not depend on details of the potential.

However, with the advent of supercomputing power in the last few years, another *ab initio* approach has emerged which makes use of a stochastic solution of the Schrodinger equation for the many-particle wave function. The methods (see, e.g., Negele, Orland, 1988) are described by names such as *Green's Function Monte Carlo and Path Integral Monte Carlo*. The approach is based on a trick: any desired property of an interacting quantum system may be represented as a functional integral of a non-interacting quantum system over a set of fictitious fields, to be discussed further below. This integral is then sampled by Monte Carlo integration methods. The central limit theorem guarantees that for sufficient samples, N, the errors will be Gaussian distributed and the variance on the integral will decrease like $O(N^{-1/2})$. To make the calculation feasible, importance sampling is used to reduce the variance. That is, in Bayesian language, the integrals are sampled according to a probability distribution which incorporates our prior knowledge of the important configurations of the fields. For example, prior knowledge may be described by an approximate wave function, such as that generated by the variational methods discussed earlier. In the Metropolis algorithm, an ergodic Markov chain is defined which samples the integral according to the thermodynamic probabilities of the fields. In the end, the result of the computer simulation is data with statistical errors much like any experiment. Maximum entropy and Bayesian methods may then be used to make inferences based on these data.

Quantum Monte Carlo and the Analytic Continuation Problem

In this section, we present a brief introduction to the quantum Monte Carlo (QMC) methods, and we motivate the actual problem solved using maximum entropy. Suppose, for example, we wish to calculate the partition function of a quantum many-body system

$$Z = Tr(e^{-\beta \hat{H}}) \,, \tag{1}$$

where $\hat{H}$ is the Hamiltonian, $\beta = 1/T$ where T is temperature, and Tr denotes the expectation value (trace) over a complete set of quantum states which describe the system. The Hamiltonian can be written as a sum of kinetic $\hat{K}$ and potential $\hat{V}$ energy terms, $\hat{H} = \hat{K} + \hat{V}$, where $\hat{K}$ and $\hat{V}$ are non-commuting operators. The quantity inside the trace in Eq. (1) is the time-evolution operator in imaginary time τ. It can be broken up into an L step Feynman path integral

$$e^{-\tau \hat{H}} = \left[e^{-\Delta\tau \hat{H}} \right]^L \simeq \left[e^{-\Delta\tau \hat{K}} e^{-\Delta\tau \hat{V}} \right]^L + O(\Delta\tau^2) \,, \tag{2}$$

where $L\Delta\tau = \beta$. The *Trotter* correction term $O(\Delta\tau^2)$ due to the non-commutativity of $\hat{K}$ and $\hat{V}$ can be ignored for sufficiently small $\Delta\tau$.

In general, the kinetic energy $\hat{K}$ is quadratic in creation and annhilation operators for the particles in the system, and the potential $\hat{V}$ is quartic in the operators for two-body interactions.

For example, for the Anderson model we will discuss later, the potential energy is the Coulomb interaction for having two electrons on the same impurity, $\hat{V} = U\hat{n}_\uparrow\hat{n}_\downarrow$, where $\hat{n}_\uparrow = \hat{d}^+_\uparrow\hat{d}_\uparrow$ is the number operator for having an up spin electron on the impurity site. We can transform exponentials of quartic operators into exponentials of quadratic operators by the analogue of completing the square. For example, consider the transformation

$$\exp\left[-\frac{\sigma^2 b^2}{2}\right] \Rightarrow \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2} - ibx\right] . \tag{3}$$

An exponential of b^2 has been transformed into an exponential of b at the expense of introducing an integration over a dummy variable, x. Analogously $\exp(-\Delta\tau\hat{V})$ for each time slice in Eq. (2) can be transformed into exponentials of $\hat{n}_\sigma$ at the expense of introducing integrations over (time-dependent) *auxiliary fields*. This is termed the *Hubbard-Stratonovich transformation*. Hamiltonians which are quadratic in creation and annhilation operators are non-interacting, and they may be solved exactly. In this fashion, the calculation of the partition function is reduced to an L dimensional integral over L auxiliary fields of a non-interacting system of particles moving in time-dependent auxiliary fields. Monte Carlo methods are then used to sample this L dimensional integral. A similar argument holds for the calculation of any other observables of the quantum system. Details of the QMC algorithms are discussed elsewhere (Loh, Gubernatis, 1990).

We consider using QMC to calculate the dynamical properties of many-electron systems such as excitation spectra, transport coefficients, magnetic susceptibility, etc. To be specific, we consider the Anderson model for dilute magnetic impurities interacting with conduction electrons in a metal. The dynamical properties can be obtained from the impurity (*retarded*) Green's function, which is simply the Fourier transform of an autocorrelation function describing the amplitude for creating an electron on an impurity at time 0 and destroying it at time t. This is written

$$G_d(\omega) = \int_0^{\infty} dt\, e^{i\omega t}\, \frac{1}{2} \sum_\sigma < [\hat{d}_\sigma(t), \hat{d}^+_\sigma(0)]_+ > = \int_{-\infty}^{\infty} d\omega\prime\, \frac{A(\omega\prime)}{\omega - \omega\prime + i\varepsilon} . \tag{4}$$

The notation is that $\hat{d}^+_\sigma(0)$ is a creation operator for putting an electron on the impurity of spin σ at time 0, <> denotes the thermodynamic expectation value, and $[,]_+$ denotes an anticommutator appropriate to Fermions. In the second part of this equation, $A(\omega)$ is termed the *spectral function*, which is a positive additive distribution function. Peaks in $A(\omega)$ correspond to the characteristic excitation energies for electrons on the impurity.

Due to the need for weighting many-particle quantum states by their thermodynamic probability, the QMC algorithm is formulated in imaginary time τ [Eqs. (1-3)]. This is fine for calculating static properties, such as the partition function, which can be formulated as integrals in imaginary time. However, we cannot calculate dynamic properties such as the real frequency quantity, $G_d(\omega)$, directly from QMC. Rather, QMC enables us to calculate an imaginary time (*Matsubara*) Green's function

$$G(\tau) = \frac{1}{2} \sum_\sigma < \hat{d}_\sigma(\tau)\hat{d}^+_\sigma(0) > . \tag{5}$$

which is defined only in the range of $0 \le \tau \le \beta$. The relationship of this quantity to the spectral function is given by the spectral representation

$$G(\tau) = \int_{-\infty}^{\infty} d\omega\, A(\omega)\, \frac{e^{-\tau\omega}}{1 + e^{-\beta\omega}} . \tag{6}$$

The determination of the real frequency spectral functions from imaginary time Green's function data is termed the *analytic continuation problem*. Fig. 1 shows typical data generated by a QMC calculation. The data are generally smooth and featureless, whereas the spectral function may have several sharp features. As discussed earlier, the data produced by QMC about $G(\tau)$ are subject to statistical errors. In the figure the errors are at the .01% level of the peak datum, so that the data appear smooth. Eq (6) is similar in many respects to a Laplace transform, particularly because a very small amount of noise on $G(\tau)$ can result in very large changes in the $A(\omega)$ obtained by trying to numerically invert this transform. The analytic continuation problem is *extremely ill-posed*.

Because of the importance of dynamical properties to all of many-body physics, there have been many attempts to solve the analytic continuation problem. Unfortunately, the developments have been a series of *ad hoc* procedures. Of course, almost all methods for solving inverse problems will work with sufficiently good data, such as a mock data set. The test should be the quality of information extracted from feasible QMC data.

The earliest attempt was to use Pade approximants (Vidberg, Serene, 1977; Hirsch, 1987). Writing the discrete frequency Fourier transform of $G(\tau)$ in the range $0 \leq \tau \leq \beta$ as $G(i\omega_m)$, with $\omega_m = m\pi/\beta\hbar$, one fits the data to the form

$$G(i\omega_m) = \frac{a_n(i\omega_m)^n + a_{n-1}(i\omega_m)^{n-1} + \ldots + a_o}{b_{n+1}(i\omega_m)^{n+1} + \ldots + b_o)}, \tag{7}$$

and then analytically continues to real frequencies. This approach ignores statistical errors. Therefore, it tends to fit the noise as well as the signal, and it can propagate noise into the resulting estimate of the spectral function. For feasible QMC data one obtains extremely noisy nonpositive images. This failure is not surprising, considering that the method does not enforce even the minimal prior knowledge about the positive and additive nature of the spectral function.

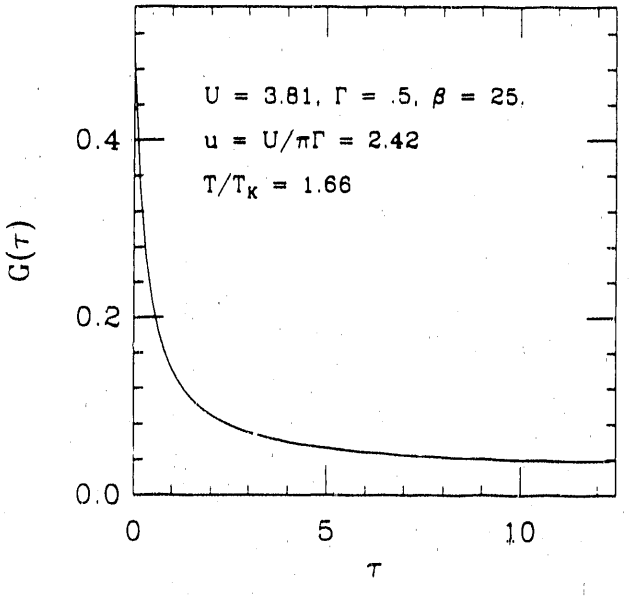

Fig. 1. A typical Matsubara Green's function, $G(\tau)$, calculated by QMC for the Anderson model with the parameters indicated.

The next approach was to use constrained least square fitting (Schuttler, Scalapino, 1985). One assumes that the QMC data are Gaussian distributed and statistically independent

$$G_{QMC}(\tau_i) = G(\tau_i) \pm \sigma_i \quad . \quad \tau_i = 0, \Delta\tau, 2\Delta\tau, \ldots, \beta \quad . \tag{8}$$

Then one functionally varies $A(\omega)$, constrained by positivity, to minimize

$$\chi^2 = \sum_i \frac{1}{\sigma_i^2} \left[G_{FIT}(\tau_i) - G_{QMC}(\tau_i) \right]^2 \quad , \tag{9}$$

where $G_{FIT}(\tau_i)$ is the $G(\tau)$ generated by a given choice of $A(\omega)$ in Eq. (6). The result of this procedure is a very small χ^2, but very noisy positive $A(\omega)$.

This was followed by constrained least squares fitting with smoothing (White, *et al.*, 1989). One minimizes

$$Q \equiv \chi^2 + b \int d\omega \left(\frac{\partial A(\omega}{\partial \omega} \right)^2 \quad . \tag{10}$$

This has an adjustable parameter, b, to control the degree of smoothing. The choice of b is essentially intuitive (*ad hoc*). This tends to produce reasonable $A(\omega)$, but there is also a tendency to produce spurious structure. In particular, there is great difficulty in accommodating both sharp and broad structure in $A(\omega)$. Nor does the method provide any error estimates, since it is not founded on probability theory. An artificial dynamics with smoothing, which has similar behavior, was also proposed (Jarrell, Biham, 1989).

The most recent attempt is the *average spectrum method* (White, 1990). One calculates

$$< A(\omega) > = \frac{\int DA(\omega) A(\omega)\, e^{-\chi^2/2}}{\int DA(\omega)\, e^{-\chi^2/2}} \quad ; \qquad A(\omega) \geq 0 \quad . \tag{11}$$

where the $DA(\omega)$ represents a Monte Carlo integration over all positive $A(\omega)$. This is a computer intensive method which produces reasonable $A(\omega)$ without smoothing. The method will also produce error estimates, i.e. simply calculate the quantity $< (A(\omega))^2 > - (< A(\omega) >)^2$. However, there is a tendency to produce spurious tails where $A(\omega)$ should be zero, and the choice of representing the image as an average spectrum violates the MaxEnt disiderata. For example, the average spectrum method will produce 1/12 for the number of blue eyed-left handed kangaroos in the Kangaroo argument (Gull, Skilling, 1984), rather than the preferred answer of 1/9.

None of these approaches regarded analytic continuation as an image reconstruction (or statistical inference) problem. The true character of the statistical errors in the QMC data was also overlooked, i.e. the QMC data are not statistically independent and Eq. (8) is invalid.

QMC Data Treatment and the Maximum Entropy Approach

A maximum entropy approach to the analytic continuation problem was first discussed by Silver, *et al.* (1990). The method was illustrated by using MEMSYS3 (Gull, Skilling, 1989) to analyze mock data in which the errors were taken to be Gaussian distributed and statistically independent. We then proceeded to analyze real QMC data for the Anderson model (Silver, *et al.*, 1990; Jarrell, *et al.*, 1990). The data were generated with a QMC algorithm (Hirsch, Fye, 1986) based on a Metropolis method. The first step in the maximum entropy approach to real QMC data is to calculate the Likelihood function.

As discussed earlier, the Metropolis procedure generates $G(\tau_i)$ data as an average over a Markov chain. Correlations in the values of the auxiliary fields are removed by binning the data over Monte Carlo times much longer than the characteristic correlation times. The binned $G(\tau_i)$ distributions become Gaussian distributed only at sufficiently large total Monte Carlo time, and

for smaller Monte Carlo times there was significant kurtosis in the $G(\tau_i)$ distributions. If N is the number of bins, and $G_n(\tau_i)$ is the average of $G(\tau_i)$ for the n'th bin, the covariance matrix for the data can be estimated from

$$C(\tau_i, \tau_j) = \frac{1}{N(N-1)} \sum_{n=1}^{N} \delta G_n(\tau_i) \delta G_n(\tau_j) , \tag{12}$$

where

$$\delta G_n(\tau) = G_n(\tau) - \frac{1}{N} \sum_{n=1}^{N} G_n(\tau) = G_{QMC}(\tau) . \tag{13}$$

We found that the covariance matrix is not diagonal, unlike the assumptions in all previous approaches to the analytic continuation problem. Rather, the covariance matrix is dense. The elements go to zero only where required by the symmetry of the problem. The eigenvalues of $C_{ij} \equiv C(\tau_i, \tau_j)$ typically spanned 4 to 6 orders of magnitude, as shown in Fig. 2.

To generalize the Likelihood function to covariant data, we first write the Green's function in vector notation:

$$(\vec{G})_i \equiv G(\tau_i) = \int_{-\infty}^{\infty} d\omega \; A(\omega) \frac{e^{-\omega \tau_i}}{1 + e^{-\beta \omega}} . \tag{14}$$

Defining

$$\delta \vec{G} = \vec{G}_{FIT} - \vec{G}_{QMC} , \tag{15}$$

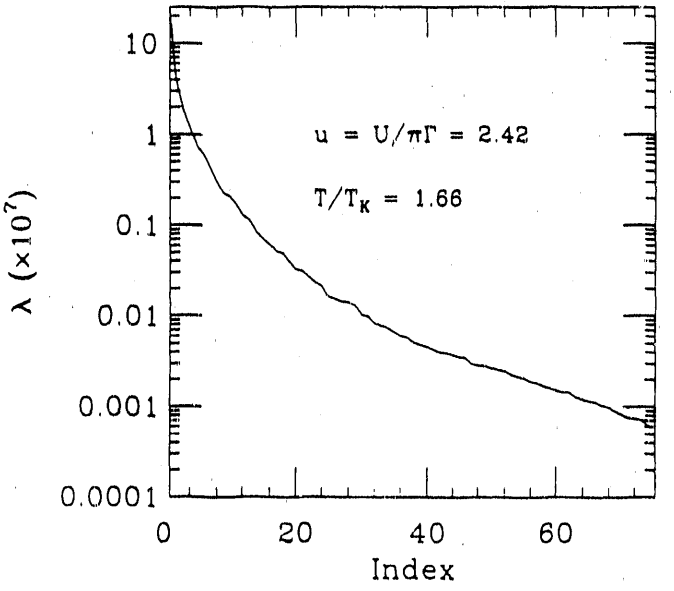

Fig. 2. Eigenvalue spectrum of the covariance matrix for the data in Fig. 1.

the generalization of the Log-Likelihood function for covariant data is (Silver, *et al.*, 1990)

$$\chi^2 \Rightarrow \delta\vec{G}^T \cdot \overleftrightarrow{C}^{-1} \cdot \delta\vec{G} \,. \tag{16}$$

However, the MEMSYS3 code requires statistically independent data. To create this, we rotate the spectral representation by an orthonormal transformation calculated by singular value decomposition of the covariance matrix, i.e.

$$\overleftrightarrow{C} = \overleftrightarrow{U} \cdot \overleftrightarrow{\Sigma} \cdot \overleftrightarrow{U}^T \,, \tag{17}$$

and

$$(\vec{D})_i = (\overleftrightarrow{U}^T \cdot \vec{G}(QMC)_i = \sum_j U^T_{ij}(\vec{G})_j \,. \tag{18}$$

Here $\overleftrightarrow{U}$ is the orthonormal matrix which diagonalizes the covariance matrix, $\vec{D}$ are the data used in the MEMSYS3 code, and $\overleftrightarrow{\Sigma}$ is a diagonal matrix whose elements are the squares of the errors on the data. We do not know of any physical significance to the rotated data space. Eq. (18) is the *image-to-data-transformation* used in the MEMSYS3 code.

Our calculation used the features of Classic Automatic MaxEnt. Although we did our best to produce Gaussian distributed data, the actual distributions showed some kurtosis. Consequently, it was important to use the automatic noise scaling which can compensate for non-Gaussian distributed data. We typically found that the errors were rescaled upwards by less than 20%. The transport coefficients of the Anderson model can be calculated by certain integrals over the spectral functions. Classic MaxEnt provides the errors on these integrals, and as we shall show this is very important for proving certain physical properties of the Anderson model.

Finally, we found that it is crucial to use an informative default model for such an ill-posed problem. We used a perturbation theory prediction (Horvatic, *et al.*, 1987) for the spectral function of the Anderson model. This choice significantly reduced the variance of the spectral functions and transport coefficients obtained. Spectral functions and transport coefficients obtained with uninformative default models had much larger errors. The perturbation theory can be expected to get the high frequency properties correctly, but it cannot be expected to correctly describe the low frequency many-body behavior. Fortunately, a Likelihood function analysis (Silver, *et al.*, 1990) of Eq. (6) shows that the QMC data provide the most information about low frequency properties. The maximum entropy procedure therefore provides a complementary combination of computer simulations with analytic theory.

The Anderson and Kondo Models

The properties of a metal with a dilute concentration of magnetic impurities has been an enduring problem in condensed matter physics. Anomalies are found experimentally in the resistivity, thermal conductivity, thermopower, specific heat and magnetic susceptibility. The resistivity anomaly is the best known: as the temperature is lowered the resistivity displays a minimum, then increases, and finally saturates as the temperature goes to zero. The qualitative understanding of this resistivity minimum is well understood in terms of the Kondo Hamiltonian (Kondo, 1964). This describes a spin one-half impurity interacting with the conduction electron spins via an antiferromagnetic coupling constant J. As the temperature is lowered toward a characteristic energy scale, T_K, the scattering rate of electrons from the impurity calculated by perturbation theory in J begins to diverge logarithmically. The increasing impurity resistivity overwhelms the decreasing phonon resistivity, and the resistivity begins to rise. The behavior as the temperature approaches zero is infinite in any finite order in perturbation theory, and in general it requires a non-perturbative treatment. The understanding which has developed from renormalization group calculations (Anderson, *et al.*, 1970; Wilson, 1975) is that the conduction electrons begin to screen the magnetic impurity. This decreases the effective impurity moment toward zero forming a local singlet composed of the

impurity spin and a compensating electron cloud. The resistivity saturates. For the spin one-half impurity, no theories explain the resistivity behavior over the entire range of interest, $T \ll T_K$, to $T \gg T_K$.

The Anderson model is a more general description of magnetic impurities in metallic alloys, which includes the possibility of charge fluctuations (changes in the number of electrons) on the impurity. The model consists of a half-filled conduction band interacting with an impurity via a hybridization matrix element V. The Hamiltonian of the (symmetric) model can be written

$$\hat{H} = \sum_{k\sigma} e_k \hat{a}^\dagger_{k\sigma} \hat{a}_{k\sigma} + \sum_{k\sigma} V(\hat{a}^\dagger_{k\sigma} \hat{d}_\sigma + \hat{d}^\dagger_\sigma \hat{a}_{k\sigma}) + U(\hat{n}_\uparrow - 1/2)(\hat{n}_\downarrow - 1/2) \ . \tag{19}$$

Here, $\hat{d}_\sigma$ is an annhilation operator for an electron on the impurity, $\hat{a}_{k\sigma}$ is an annhilation operator for a conduction electron of energy e_k and spin σ, and $\hat{n}_\sigma = \hat{d}^\dagger_\sigma \hat{d}_\sigma$ is the number operator for a spin σ electron on the impurity site. There is a Coulomb energy U for two electrons to be on the same impurity. In the limit of large U charge fluctuations on the impurity are suppressed, and the Anderson model should reduce to the Kondo model.

To develop our expectations for the spectral function, we consider some limits. First, if the hybridization is taken to zero, the energy levels of the impurity are $E = 0$ for one electron on the impurity and $E = U/2$ for zero or two electrons on the impurity. We, therefore, expect single-electron excitations at $\omega = \pm\ U/2$. If now we turn on the hybridization, we expect these excitations to be Lorentzian broadened by a hybridization width $\Gamma = N(0)\pi V^2$, where $N(0)$ is the conduction band density of states at the Fermi level derived from e_k. The limit of large U provides another, much smaller, energy scale for single particle excitations. Then, a transformation (Schrieffer, Wolff, 1966) of the Anderson model yields the Kondo Hamiltonian with antiferromagnetic coupling constant, $J = -8\Gamma/\pi N(0)U$. Using high temperature perturbation theory (Haldane, 1978) in J coupled with numerical results (Jarrell, *et al.*, 1990), the Kondo temperature is given by $T_K \approx 0.515(1+1/2u)\Gamma\sqrt{u}$ $\exp(-\pi^2 u/8)$, where $u \equiv U/\pi\Gamma$. By the particle-hole symmetry of the symmetric model we expect a *Kondo peak* in the spectral function centered at $\omega = 0$ of width comparable to T_K. To discuss scaling properties, we first remove an arbitrary energy scale by dividing all energies by Γ, and we discuss a *scaled spectral function* $\pi\Gamma A(\omega)$. If U is sufficiently large compared to T such that there are no significant charge fluctuations on the impurity site, the only energy scale at low frequencies is the Kondo temperature, T_K. We expect the scaled spectral function to be a universal function of ω/T_K and T/T_K for ω small compared with U. We refer to a lack of dependence on U as *universality*.

In the dilute impurity limit for the Anderson model the single-particle spectral function may be related directly to the transport coefficients (Bickers, *et al.*, 1987). The resistivity is given by

$$\frac{\varrho(0)}{\varrho(T)} = -\frac{1}{\pi\Gamma} \int_{-\infty}^{\infty} d\omega \, \frac{\partial f}{\partial \omega} \, A^{-1}(\omega) \ , \tag{20}$$

where $f(\omega)$ is the Fermi function, $1/(1 + \exp(-\beta\omega))$. At sufficiently low temperatures the Fermi factor, $\partial f/\partial\omega$, in the resistivity ensures that only the low frequency part of the scaled spectral function contributes. Therefore the transport coefficients can also be expected to be universal functions of T/T_K. At sufficiently high temperatures, the Fermi factor will mix in the high frequency peaks at $\omega = \pm\ U/2$ resulting in non-universal behavior.

Horvatic *et al.* have developed a perturbation expansion for the spectral function with u as the small parameter. While their expansion is in principle absolutely convergent, in practice only a few diagrams in the expansion can be calculated. Their results should be valid at small u and become increasingly poor at large u. While they recover the main features of the spectral function discussed above, the $\omega = 0$ peak in their spectral function is not a universal functi nor is the resistivity calculated using their spectral function in Eq. (20). A critical question we wish to address is how high must u be in order to obtain universality.

The Kondo limit at large u leads to several expectations for the spectral functions and transport coefficients. The detailed shape of the spectral function at zero temperature has been predicted to be (Doniach, Sunjic, 1970)

$$\pi\Gamma A(\omega)_{T=0} = Re\sqrt{\frac{i\Gamma_K}{\omega+i\Gamma_K}}, \tag{21}$$

although they did not predict the value of Γ_K. Wilson (1975) carried out a famous numerical renormalization group study of the Kondo model. He defined a low temperature energy scale, T_K^o, by $T_K^o\chi(0) = {\mu_B}^2$, where $\chi(T)$ is the magnetic susceptibility. He found that $T_K^o = 2.4\ T_K$, where $1/2.4$ is refered to as the *Wilson number*. Based on Wilson's study, Nozieres (1974) predicted that at low temperatures the resistivity would show Fermi liquid behavior with

$$\lim_{T \ll T_K} \frac{\varrho(T)}{\varrho(0)} \simeq 1 - \alpha\left(\frac{T}{T_K}\right)^2. \tag{22}$$

(Note, that α is different from the statistical regularization parameter in MaxEnt.) Nozieres suggested that α is determined by the Wilson number to have the value 1.03. At temperatures high enough to apply perturbation theory in J, the resistivity is expected to be given by (Hamann, 1967; Nagaoka, 1965; Suhl, 1965)

$$\text{for } T \gtrsim T_K^N \quad \frac{\varrho(T)}{\varrho(0)} \simeq \frac{1}{2}\left[1 - \frac{\ln T/T_K^N}{\sqrt{\ln^2 T/T_K^N + 3\pi^2/4}}\right]. \tag{23}$$

Here T_K^N is the "Nagaoka" Kondo temperature defined by where the resistivity goes through one-half its peak value. One goal of our calculations is to test these predictions and to determine the parameters Γ_K, α, and T_K^N along with their relation to T_K. In addition, we wish to obtain the universal resistivity over the entire range from $T \ll T_K$, to $T \gg T_K$.

Results

Figure 3 shows typical results (Silver, *et al.*, 1990) for the scaled spectral function $\pi\Gamma A(\omega)$. The spectral function obtained from the QMC data using maximum entropy is labeled QMC-ME. The perturbation theory prediction of Harvatic *et al.*, is labeled H. The ratio of the two results is labeled QMC-ME/H. While the H prediction has the qualitatively correct structure, the QMC-ME Kondo peak is depressed and broadened compared to H and the $U/2$ peaks are slightly enhanced. This preserves the sum rule on the spectral function. In general, we find that QMC-ME and H agree for $u < 1.2$. However, with increasing u the QMC-ME Kondo peaks becomes increasingly depressed and broadened compared to H, and the QMC-ME $U/2$ peaks become increasingly enhanced compared to H.

In Fig. 4 the QMC-ME spectral functions are plotted (Silver, *et al.*, 1990) against ω/T_K at fixed $T/T_K = 1.5$ and for a variety of u. The semilog scale emphasizes the behavior at low frequencies. One can see that the QMC-ME spectral density is approximately universal (independent of u) for $\omega/T_K \lesssim 20$, although the high frequency behavior around the $U/2$ peak is non-universal. While error estimates at individual ω/T_K points are very large, the error estimates on integrals of the spectral function are small and meaningful. The inset in Fig. 4 shows that the average of $\pi\Gamma A(\omega)$ for $\omega/T_K \lesssim 20$ is universal for $u > 1.25$ within statistical error. In contrast, the H predictions are distinctly non-universal in the same frequency range. Only for $u \lesssim 1.2$ where both H and QMC-ME agree within statistical errors is QMC-ME non-universal.

Figure 5 shows (Silver, *et al.*, 1990) the evolution of the universal Kondo peaks as a function of T/T_K at fixed u 2.0. At $T/T_K \gtrsim 10$ the Kondo peak is absent. With decreasing T/T_K the Kondo peak grows and narrows until at $T/T_K \lesssim 0.2$ the height of the central peak has almost approached

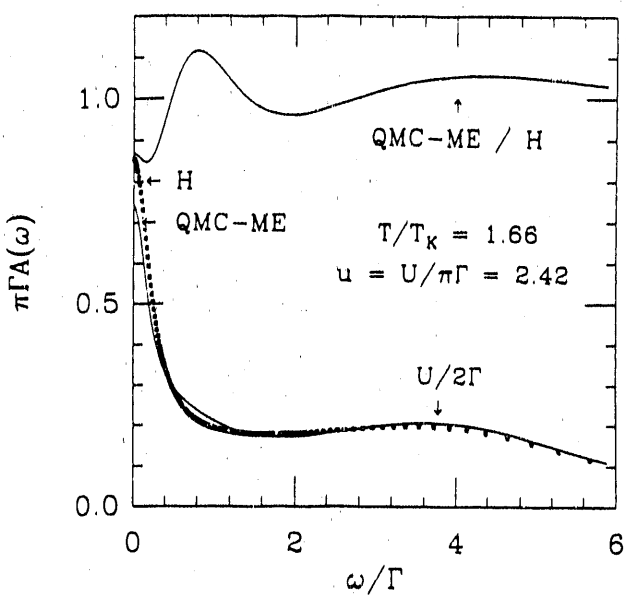

Fig. 3. QMC-ME labels the quantum Monte Carlo and maximum-entropy results for the spectral density of the symmetric Anderson model for the parameters indicated. H is the prediction of the self-consistent second-order perturbation theory of Horvatic *et al.* for the self-energy. QMC-ME/H labels the ratio of the two calculations.

the Freidel sum rule value of 1.0. The inset in Fig. 5 shows the screened local moment calculated by QMC, $T\chi(T)/g^2$, plotted versus T/T_K for $u = 2.0$. The growth of the Kondo peak with decreasing T/T_K is correlated with the screening of the impurity local moment by the conduction electrons. Plotted as the dashed line labeled DS in this figure is the Doniach-Sunjic prediction for the shape of the Kondo peak at zero temperature. We find that a best fit to the $T/T_K = 0.2$ QMC-ME curve is obtained with $\Gamma_K = 2.5\ T_K$, which is remarkably the inverse of the Wilson number.

Figure 6 shows (Jarrell, *et al.*, 1990) the resistivity ratio, $\varrho(T)/\varrho\ (0)$, plotted vs. u for $T/T_K = 1.5$. Universality is demonstrated by the fact that for $u > 1.5$ the value of the resistivity saturates to a constant, whereas the Horvatic *et al.* perturbation theory prediction continues to rise monotonically. We wish to emphasize that the error bars on our resistivity come from our estimates of the covariance matrix of the QMC calculation, and the errors are propagated by our classic MaxEnt code into integrals over the spectral function. The increasing error bars with increasing u are due to several factors. The QMC statistical errors become larger with increasing u because the representation we are expanding about becomes worse. The perturbation theory, when used as a MaxEnt default model, diverges from the QMC data with increasing u. The error estimates are much larger when uninformative default models are used, making it more difficult to prove universality.

In Fig. 7 the resistivities are plotted versus $ln(T/T_K)$ for several values of u (Jarrell, *et al.*, 1990). One can see that these define a universal resistivity curve (independent of u) over the desired range from $T \ll T_K$ to $T \gg T_K$. Since at $T \ll T_K$ the resistivity ratio is very close to one, it is difficult to determine reliably the difference between the resistivity ratio and one. The dotted line in Fig. 7 is the Fermi liquid result, Eq. (22), with α an adjustable parameter. We found a best fit for $(T/T_K)^2 < 0.1$ of $\alpha = 0.83 \pm 0.06$. The quoted error reflects only statistical sources of error, but there is also a potential for systematic error. A subtle point of the maximum entropy method is a built-in bias toward the default model of approximately one standard deviation. The default

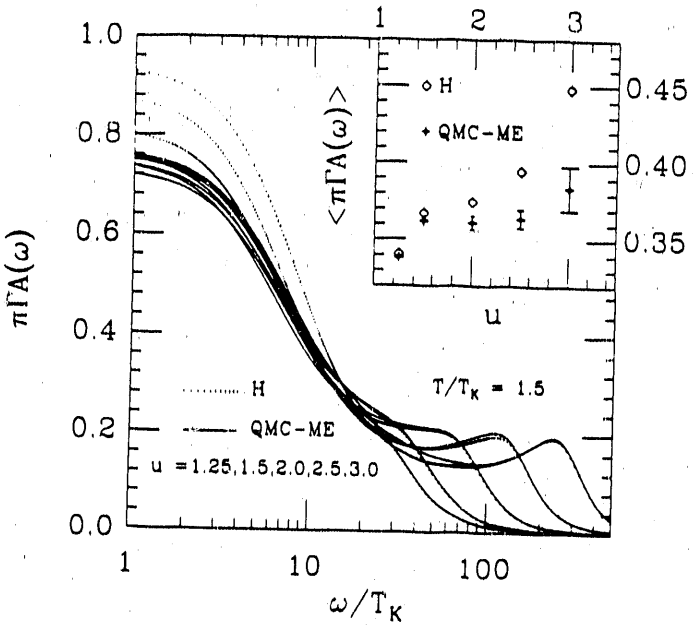

Fig. 4 Study of the Kondo scaling of the spectral density at fixed $T/T_K = 1.5$ and for the values of $u \equiv U/\pi\ \Gamma$ indicated. QMC-ME labels the Monte Carlo and maximum entropy results. They are universal functions of ω/T_K (independent of u) for ω/T_K less than 20, and they are nonuniversal for larger ω/T_K in the vicinity of the ω = U/2 peaks. Inset: Average of $\pi\ \Gamma A(\omega)$ for $\omega/T_K < 20$ plotted vs. u. H labels the predictions of Horvatic *et al.* for the same values of u, with higher $\pi\ \Gamma A(0)$ corresponding to larger u. The H theory is distinctly non-universal.

model was the Horvatic *et al.* perturbation theory which had a larger resistivity ratio than the QMC-ME calculation (Alternatively, it had a smaller α, although the Horvatic *et al.* calculation does not show Kondo universality.) For $(T/T_K)^2$ <0.1 the Nozieres prediction is approximately one standard deviation above the QMC-ME results for each T/T_K value.

The solid line in Fig. 7 is the Hamann *et al.* prediction, Eq. (23), obtained by fitting the data for $T > T_K$ by adjusting T_K^N so that the curves agree when $\varrho(T_K^N)/\varrho\ (0) = 0.5$. We find T_K^N/T_K = 2.5, which is remarkably again the inverse of the Wilson number for the Kondo problem. Our results agree well with Eq. (23) at high temperatures, but they diverge as expected for $T < T_K$. However, for the highest temperatures simulated the resistivity shows some slight non-universality and the resistivity ratio falls below the Hamann result. In this limit the Anderson model moment is thermally reduced due to occupation of the zero electron and two electron states on the impurity, whereas the moment in Hamann's calculation is fixed at one-half. For finite u and $T \gg T_K$ the impurity moment which scatters the electrons is reduced and so is the resistivity.

Conclusions and Discussion

We have verified all of our expectations about the dynamical properties of the symmetric Anderson model. For $u < 1.2$ we obtain universal Kondo scaling for low frequency properties including the Kondo peak in the spectral function and the resistivity. We find agreement with the Doniach-Sunjic expression for the zero-temperature shape of the spectral function and with the Hamann-Nagoaka-Suhl expression for the resistivity for $T > T_K$. Our results are consistent with Nozieres prediction

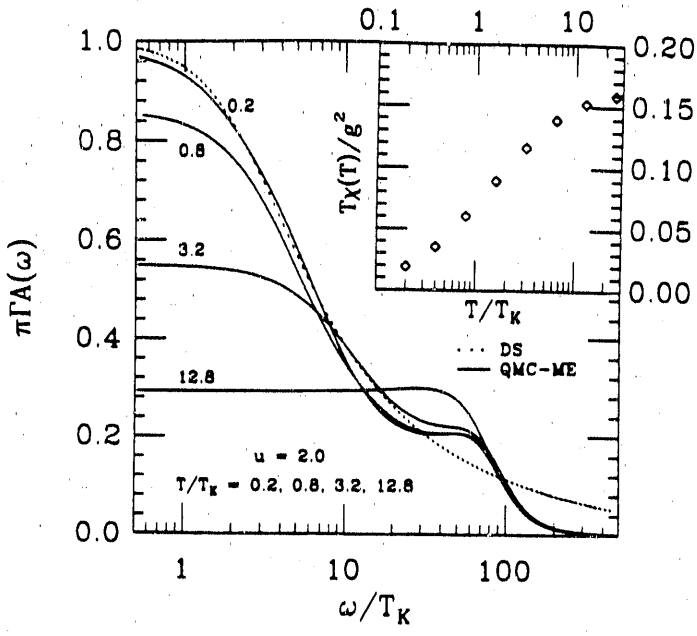

Fig. 5. QMC-ME labels the quantum Monte Carlo and maximum entropy results for fixed $u = 2.0$ and varying T/T_K as indicated. The Kondo peak is approximately independent of u for $\omega/T_K < 20$. DS labels the Doniach-Sunjic expression, Eq. (21), for the zero temperature shape of the Kondo peak. We find a best fit for $\Gamma_K = 2.5\ T_K$. Inset: Screened local moment, $T\chi(T)/g^2$, plotted vs. T/T_K.

of Fermi liquid behavior for the resistivity at $T \ll T_K$, although the coefficient α of the $(T/T_K)^2$ term may be slightly smaller than Nozieres value. Moreover, we find a remarkable convergence for the numerical parameters of these theories, summarized in the following equation,

$$T_K^N \ = \ \Gamma_K \ = \ T_K^o \ = \ 2.5 T_K \ . \tag{24}$$

Here, T_K^N is the Nagaoka Kondo temperature defined by where the resistivity ratio goes through one-half; Γ_K is the width in the Doniach-Sunjic expression for the zero temperature spectral function; T_K^o is Wilson's Kondo temperature determined by numerical renormalization group calculations; 2.5 is approximately the inverse of the famous Wilson number; T_K is the Kondo temperature determined by high temperature perturbation theory in the impurity-conduction electron spin coupling constant J. The remarkable simplicity and convergence between the parameters of these theories certainly deserves an explanation.

Perhaps most significantly for the long term, we have established a general methodology (Silver, *et al.*, 1990) for the *ab initio* calculation of the dynamical properties of many-body systems. This combines quantum Monte Carlo to generate Matsubara Green's functions in imaginary time, the maximum entropy method to solve the extremely ill-posed problem of numerically inverting the spectral representation, and perturbation theory and other forms of prior knowledge to provide informative default models needed to minimize the variance of the maximum entropy results. Our method provides for full error propagation from the raw QMC data all the way to integrated properties of the spectral function, such as the transport coefficients. Moreover, the maximum entropy method provides a natural combination of analytic theory to obtain high frequency properties with quantum Monte Carlo to obtain low frequency (many-body) properties.

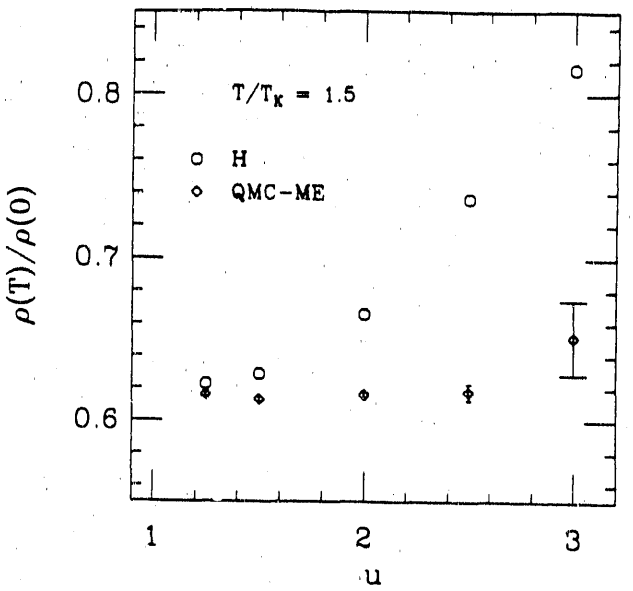

Fig. 6. The resistivity ratio $\varrho(T)/\varrho\,(0)$ plotted vs. $u \equiv U/\pi\Gamma$ when $T/T_K = 1.5$. The open circles are from the perturbation theory of Horvatic *et al.* For $u < 1.2$ there was no significant difference between the perturbation theory and the Monte Carlo results. Universality of the Monte Carlo data is indicated by the fact that for $u > 1.5$ the value of the ratio saturates to a constant within MaxEnt statistical errors, whereas the perturbation theory continues to rise monotonically.

Our method has already been applied to several other problems in computational condensed matter physics. These include verifying the Haldane conjecture about Heisenberg spin chains (Deisz, *et al.*, 1990), calculating gap states in superconductors induced by magnetic impurities (Jarrell, *et al.*, 1990) and determining the density of states of the Hubbard models (Silver, White, 1990) which are perhaps relevant to high T_c superconductivity. We are currently applying the method to the calculation of the dynamical magnetic susceptibility of the Anderson model, to the asymmetric Anderson model to better compare with experimental data, and to the X-ray edge singularity problem.

We would like to emphasize a number of technical features of the application of MEMSYS3 to QMC. The analytic continuation of QMC data is an extremely ill-posed problem. The number of good measurements defined by Classic MaxEnt is typically 5-8. The posterior distribution of the statistical regularization parameter in MaxEnt, α, is not sharply peaked. While in principle we should have marginalized over the distribution of α, we have not yet done so. Recently a new MaxEnt algorithm has been proposed (Bryan, 1990) which marginalizes over α for such over-sampled inverse problems. Nor have we as yet implemented pre-blur (*intrinsic correlation functions*), as is currently suggested to overcome the tendency of classic MaxEnt to overfit data (Skilling, 1990). Rather, we have minimized the variance of the MaxEnt images and integrated quantities by relying on perturbation theory for the Anderson model to provide an informative default model. For other QMC applications such high quality default models may not exist. Sometimes prior knowledge of sum rules and moments may be used to construct a reasonable default model according to the principle of maximum entropy. However, for most future QMC problems we expect marginalizing over α, and perhaps implementing pre-blur, to become more important.

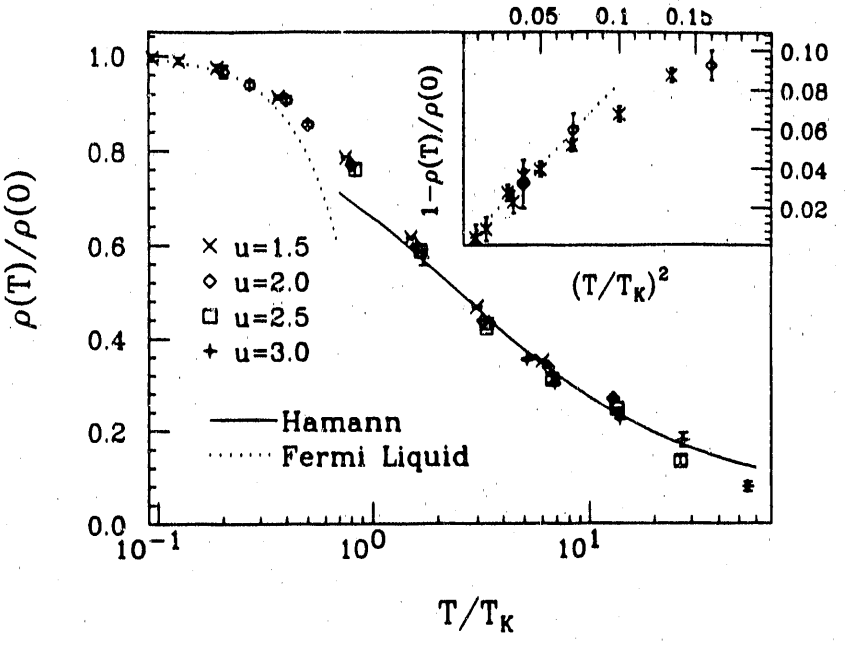

Fig. 7. $\varrho(T)/\varrho\ (0)$ vs. T/T_K at different values of u. The solid line is from the self-consistent high temperature calculation of Hamann *et al.*, Eq. (23). It is fit to our results by varying T_K^N so that the curves agree when $\varrho(T_K^N)/\varrho\ (0) = 0.5$. We find $T_K^N = 2.5\ T_K$. As expected, this prediction fails at low temperatures. The dotted line is the Fermi liquid result due to Nozieres, Eq. (22), with a best fit of $\alpha = 0.83 \pm 0.06$. The resistivity is roughly logarithmic for $0.5 < T/T_K < 4.0$.

Classic MaxEnt error estimates on integrated quantities were essential to the proof of Kondo universality. We have found these estimates to be reliable. Repeating the QMC calculation with a different random number seed produced resistivities within one standard deviation of equivalent runs. A longer QMC run reduced the error, and the new resistivity was inside the error estimates of the shorter QMC run. Automatic noise scaling was also essential to our success. Without it, the Classic MaxEnt spectral functions are ringing and clearly overfitting. We found the noise scaling was typically less than 20%, or it could be reduced to less than 20% by a longer QMC run. Presumably, the longer runs produce data which are more Gaussian distributed. Because of the 4-6 orders of magnitude range of the eigenvalues of the covariance matrix, we had to report the QMC data to 6-7 digits and to run the MEMSYS3 code in double precision. We found that failure to do so would result in large noise scalings. The data could be processed using MEMSYS3 in 5-10 minutes of Sun Sparcstation time, whereas generating the raw QMC data took from 1 to 4 hours on a CRAY Y-MP for each set of Anderson model parameters. This CRAY time is, nevertheless, orders of magnitude smaller than would be required by direct methods of analytic continuation, such as simple Padé approximants. The use of maximum entropy and Bayesian methods for data analysis can dramatically reduce the cpu time required for computer simulations.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Depart. of Energy, Basic Energy Sciences, Division of Materials Research. We thank the Ohio Supercomputer Center for a grant of computer time. We gratefully acknowledge useful discussions with D. L. Cox, J. Deisz, F. D. M. Haldane, C. Jayaprakash, and J. W. Wilkins.

REFERENCES

Anderson, P. W. (1961), "Localized Magnetic States in Metals," Phys. Rev. **124**, 41.

Anderson, P. W., Yuval, G., Hamann, D. R. (1970), "Exact Results in the Kondo Problem. II. Scaling Theory, Qualitatively Correct Solution, and Some New Results on One-Dimensional Classical Statistical Models," Phys. Rev. B**1**, 4664.

Bickers, N. E., Cox, D. L., Wilkins, J. W. (1987), "Self-consistent large-N expansion for normal state properties of dilute magnetic alloys," Phys. Rev. B**36**, 2036.

Bryan, R. K. (1990), "Maximum entropy analysis of oversampled data problems," Eur. Biophys. J. **18**, 165-174.

Deisz, J., Jarrell, M., Cox, D. L. (1990), "S(Q,ω) for the S=1/2 and S=1 1-D Heisenberg Antiferromagnet: A Quantum Monte Carlo Study," to appear.

Doniach, S., Sunjic, M. (1970), "Many-electron singularity in X-ray photoemission and X-ray line spectra from metals," J. Phys. C**3**, 285.

Gull, S. F., Skilling, J., (1984), "Maximum Entropy Method in Image Processing," IEE Proc. **131**(F), 646.

Gull, S. F. (1989), "Developments in Maximum Entropy Data Analysis," in *Maximum Entropy and Bayesian Methods*, ed. J. Skilling (Kluwer Academic, Dordrecht).

Gull, S. F., Skilling, J. (1989), "Quantified Maximum Entropy, MEMSYS3," produced by Maximum Entropy Data Consultants, Ltd.

Haldane, F. D. M. (1978), "Theory of the atomic limit of the Anderson model: I. Perturbation expansions re-examined," J. Phys. C**11**, 5015.

Hamann, D. R. (1967), "New Solution for Exchange Scattering in Dilute Alloys," Phys. Rev. **158**, 570.

Hirsch, J. E., Fye, R. M. (1986), "Monte Carlo Methods for Magnetic Impurities in Metals," Phys. Rev. Lett. **56**, 2521.

Hirsch, J., (1987), "Simulation of Magnetic Impurities in Metals," in *Quantum Monte Carlo Methods*, ed. M. Suzuki (Springer-Verlag, Heidelberg), p. 205.

Horvatic, B., Sokcevic, D., Zlatic, V. (1987), "Finite-Temperature spectral density for the Anderson Model," Phys. Rev. B**36**, 675.

Jarrell, M., Biham, O. (1989), "Dynamical Approach to Analytic Continuation of Quantum Monte Carlo Data," Phys. Rev. Lett. **63**, 2504.

Jarrell, M., Gubernatis, J. E., Silver, R. N., Sivia, D. S. (1990), "The Transport Coefficients of Dilute Magnetic Alloys," Phys. Rev. Rap. Comm., to appear.

Jarrell, M., Sivia, D. S., Patton, B. R. (1990), "Gap States in Dilute Magnetic Alloy Superconductors: A Quantum Monte Carlo Study," Phys. Rev. Rapid. Comm., to appear.

Kondo, J. (1964), "Resistance Minimum in Dilute Magnetic Alloys," Prog. Theor. Phys. **32**, 37.

Loh, E., Gubernatis, J. E., (1990), "Stable Numerical Simulations of Models of Interacting Electrons in Condensed-Matter Physics," in *Electronic Phase Transitions*, eds. W. Hanke, Y. Kopaev (Elsevier, New York).

Nagaoka, Y. (1965), "Self-Consistent Treatment of Kondo's Effect in Dilute Alloys," Phys. Rev. **138**, A1112.

Negele, J. W., Orland, H. (1988), *Quantum Many-Particle Systems*, Addison Wesley, Ch. 8.

Nozieres, P. (1974), "A Fermi-liquid Description of the Kondo problem at Low Temperatures," J. Low Temp. Phys. **17**, 31.

Schrieffer, J. R., Wolff, P. A. (1966), "Relation between the Anderson and Kondo Hamiltonians," Phys. Rev. **149**, 491.

Schuttler, H. B., Scalapino, D. J. (1985), "Monte Carlo Studies of the Dynamics of Quantum Many-Body Systems," Phys. Rev. Lett. **55**, 1204; Phys. Rev. B**34**, 4744 (1986).

Silver, R. N., Sivia, D. S., Gubernatis, J. E. (1990), "Dynamical Properties from Quantum Monte Carlo by the Maximum Entropy Method," in *Quantum Simulations of Condensed Matter Phenomena*, eds. J. D. Doll, J. E. Gubernatis (World Scientific, Singapore), p. 340.

Silver, R. N., Sivia, D. S., Gubernatis, J. E. (1990), "Maximum Entropy Method for Analytic Continuation of Quantum Monte Carlo Data," Phys. Rev. B**41**, 2380.

Silver, R. N., Sivia, D. S., Gubernatis, J. E., Jarrell, M. (1990), "Spectral Densities of the Symmetric Anderson Model," Phys. Rev. Lett. **65**, 496.

Silver, R. N., White, S. (1990), in progress.

Skilling, J. (1990), this conference, "The MaxEnt Movie."

Suhl, H. (1965), "Dispersion Theory of the Kondo Effect," Phys. Rev. **138**, A515.

Vidberg, H. J., Serene, J. W. (1977), "Solving the Eliashberg Equations by Means of N-Point Pade Approximants," J. Low Temp. Phys. **29**, 179.

White, S. R., Scalapino, D. J., Sugar, R. L., Bickers, N. E. (1989), "Monte Carlo Calculations of Dynamical Properties of the Two-Dimensional Hubbard Model," Phys. Rev. Lett. **63**, 1523.

White, S. R. (1990), submitted for publication.

Wilson, K. G. (1975), "The renormalization group: Critical phenomena and the Kondo problem," Rev. Mod. Phys. **47**, 773.

END

DATE FILMED

10/29/90