NOTICE

CERTAIN DATA CONTAINED IN THIS DOCUMENT MAY BE DIFFICULT TO READ IN MICROFICHE PRODUCTS.

LA-UR-93-1296

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

TITLE: QUANTUM MONTE CARLO BY MESSAGE PASSING

AUTHOR(S): J. Bonca, T-11
J. E. Gubernatis, T-11

SUBMITTED TO: For the proceedings of the Workshop "Concurrent Computing in the Physical Sciences," held at Baten Rouge, Louisiana, Feb. 18-20, 1993. To be published by Springer-Verlag, Heidelberg.

By acceptance of this article, the publisher recognized that the U.S. Government retains a nonexclusive, royalty free license to publish or reproduce the published form of this contribution or to allow others to do so for U.S. Government purposes.

The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

MASTER

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

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

Quantum Monte Carlo by Message Passing

J. Bonča and J. E. Gubernatis
Theoretical Division and Center for Nonlinear Studies
Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.

Abstract. We summarise results of quantum Monte Carlo simulations of the degenerate single-impurity Anderson model using the impurity algorithm of Hirsch and Fye. Using methods of Bayesian statistical inference, coupled with the principle of maximum entropy, we extracted the single-particle spectral density from the imaginary-time Green's function. The variations of resulting spectral densities with model parameters agree qualitatively with the spectral densities predicted by NCA calculations. All the simulations were performed on a cluster of 16 IBM R6000/560 workstations under the control of the message-passing software PVM. We described the trivial parallelisation of our quantum Monte Carlo code both for the cluster and the CM-5 computer. Other issues for effective parallelisation of the impurity algorithm are also discussed.

1. Introduction

In spite of the fact that the single-impurity Anderson model [1] was first proposed 30 years ago as a model for the properties of dilute magnetic alloys, theoretical and numerical work on the model remains very active because it is one of the simplest paradigms for a system of strongly interacting electrons. Over the years, considerable progress has been made in understanding the properties of the model by several significant advances in analytic and numerical techniques [2, 3]. These techniques have sought to calculate various static and dynamic correlation functions to reveal the relevance of the model for such many-body phenomena as the Kondo effect, mixed valence fluctuations, and local magnetic moment formation that are observed in dilute magnetic alloys.

Although the main features of the spectral density function of the degenerate model, such as the position of the broad side peaks and the existence of a sharp resonance close to the Fermi energy, are likely well reproduced by different analytical and numerical methods [4, 5, 6, 7, 8], relative spectral weights

---

* Permanent address: J. Stefan Institute, University of Ljubljana, 61111 Ljubljana, Slovenia.

and their temperature dependence often seem to be dependent on the underlying approximation. Therefore, there is a need for a method which calculates the spectral density function of the degenerate Anderson model at arbitrary interaction strength U, hybridisation Γ, degeneracy N, and temperature T. In many respects, the quantum Monte Carlo method fulfills this need. Incorporating higher degeneracies into the model is also needed. One the one hand, it would enable a comparison of quantum Monte Carlo results with the non-crossing approximation (NCA) [7] and the numerical renormalisation group (NRG) calculations [4, 5, 6] and, on the other hand, it would bring the calculations closer to physical systems. For example, a degeneracy of $N = 6$ matches the degeneracy of a Ce impurity in a host with cubic symmetry and strong spin-orbit splitting.

In spite of numerous experimental and theoretical works in the field of dilute magnetic alloys, some disagreement still exists between theory and experiment, and even among different experimental groups [8, 9]. Spectroscopic investigations are primarily done on ordered alloys and compounds, with the assumption that the single impurity theory can be directly applied to these concentrated systems. The generally accepted belief is that the single impurity model reproduces the main spectral features in Ce or Yb heavy fermion compounds [8]; however, recent studies by Joyce *et al.* [9] show features near $\omega = 0$ that do not appear to scale with the Kondo temperature T_K or display the appropriate temperature dependence. These findings remain a puzzle.

In this paper, we give some examples of our recent calculations of the properties of the degenerate Anderson model [10], obtained by means of the quantum Monte Carlo method of Hirsch and Fye [11], which we extended to higher degeneracies. We will discuss the sources of computational intensity in this algorithm and the embarassingly simple way we used a cluster of workstations to reduce the computation time. We will also discuss other issues of parallelisation to reduce the computation time even further.

2. Model and Methods

We treated the following form of the degenerate Anderson model [7]

$$\begin{aligned} H &= H_0 + H_1 \\ H_0 &= \sum_{km} \epsilon_k n_{km} + \sum_{km} V_{km}(c^\dagger_{km} f_m + f^\dagger_m c_{km}) + \sum_m \epsilon_m n_m \\ H_1 &= \frac{1}{2} \sum_{m,m'} U_{mm'} n_m n_{m'} \end{aligned} \tag{1}$$

where $c^\dagger_{km}$ creates a state in the conduction band with the energy ϵ_k in the channel m, $f^\dagger_m$ creates an orbital state m at the impurity site with the un

renormalised energy ϵ_m, and n_{km} and n_m are the numbers operators for the conduction band and orbitals at the impurity site. V_{km} represents hybridisation between the conduction band and the localised impurity states. We will assume that the conduction band is infinitely wide and structureless; therefore, V_{km} is neither energy nor channel dependent. This assumption leads to the simple relation for the impurity level half-width $\Gamma = \pi N(0)V$, where N(0) is the energy density of states per spin at the Fermi energy. The symmetric matrix $U_{mm'}$, with the additional condition $U_{mm} = 0$, represents the Coulomb repulsion between two electrons occupying different orbitals at the impurity site. Furthermore, we associate the channel index m with the magnetic quantum number $m \equiv m_j$ since we want to model systems with strong spin-orbit coupling, such as Ce impurities in a metal. In particular, the low-lying multiplet in Ce has a total angular momentum $j = 5/2$ and therefore a degeneracy $N = 2j + 1 = 6$, which represents the highest degeneracy reached in our calculations.

We will be mainly concerned with the computation of the single-particle spectral density associated with the impurity state. Several important features of this function are known quite generally. The imaginary-time Green's function $G(\tau)$, which we will obtain using quantum Monte Carlo simulation procedures, is directly connected to the spectral function $A(\omega)$ through the following relation [12]

$$G(\tau) = \int_{-\infty}^{+\infty} d\omega \frac{e^{-\tau\omega}}{1 + e^{-\beta\omega}} A(\omega) \tag{2}$$

where β is the inverse temperature. In the case of the particle-hole symmetry, the Green's function obeys the relation $G(\tau) = G(\beta - \tau)$ and therefore $A(\omega)$ is an even function of frequency. Furthermore, $A(\omega)$ obeys the following relations [12]

$$\int_{-\infty}^{+\infty} d\omega A(\omega) = 1 \tag{3}$$

$$A(\omega) \geq 0 \tag{4}$$

These relations allow one to interpret the spectral density as a probability function.

Using Eq. (2) to determine the spectral function is also known as the analytic continuation problem because the dynamics defined on imaginary-time axis is used to determine real-time dynamics. There are several difficulties associated with this problem: the Green's function is almost insensitive to changes in $A(\omega)$ at large frequencies due to the exponentially small kernel. This insensitivity makes the problem extremely ill-posed; that is, small variations in the values of $G(\tau)$ can lead to major variations in the solution $A(\omega)$. With $G(\tau)$ being determined by a Monte Carlo procedure, variations in the data

(noise) are a fact. Furthermore, the number of data is smaller than the desired number of values of $A(\omega)$; thus we cannot solve the problem exactly. Out of the infinite number of possible solutions we will use the most probable one [13]. The notion of probability enters the problem through Bayes's theorem.

Given two events a and b, Bayes's theorem says [14]

$$\Pr[a, b] = \Pr[a|b] \Pr[b] = \Pr[b|a] \Pr[a] \tag{5}$$

where $\Pr[a]$ is the probability of a and $\Pr[a|b]$ is the conditional probability of a given b, and $\Pr[a, b]$ is the joint probability function for a and b. Our events are the functions $\tilde{G}(\tau)$ and $A(\omega)$, where $\tilde{G}(\tau)$ is our estimated (measured) value of $G(\tau)$. Our *initial* criterion for a best solution will be the function $A(\omega)$ that maximises $\Pr[A|\tilde{G}]$. We take

$$\Pr[A|\tilde{G}] = \Pr[\tilde{G}|A] \Pr[A] / \Pr[\tilde{G}] \tag{6}$$

$\Pr[A|\tilde{G}]$ is called the posterior probability, $\Pr[\tilde{G}|A]$, the likelihood function, and $\Pr[A]$, the prior probability. $\Pr[\tilde{G}]$ is a normalisation constant [15], called the evidence.

Equation (6) transfers the problem of specifying the posterior probability to the problem of specifying the likelihood function and the prior probability. These latter functions are generally ones about which we can either make reasonable assumptions or have specific knowledge.

The principle of maximum entropy uses the *a priori* knowledge that the spectral density satisfies (3) and (4), and it enters the process by specifying the prior probability of the spectral density, namely

$$\Pr[A] \propto e^{-\alpha S} \tag{7}$$

where

$$S = \sum_i [A_i - m_i - A_i \ln(A_i/m_i)] \tag{8}$$

The function m_i is called the default.

Our results are calculated from [16, 17]

$$\langle A \rangle = \int d\alpha \Pr[\alpha|\tilde{G}, m] \hat{A}(\alpha) \tag{9}$$

where the conditional probability function $\Pr[\alpha|\tilde{G}, m]$ is found by using Bayes's theorem. Besides (7), an important ingredient is the choice of the likelihood function

$$\Pr[\tilde{G}|A] \propto e^{-\chi^2/2} \tag{10}$$

where χ^2 is the least-squares function

$$\chi^2 = \frac{1}{2}\sum_{ik}(\bar{G}_i - \sum_j K_{ij}A_j)\,C_{ik}^{-1}\,(\bar{G}_k - \sum_j K_{kj}A_j) \qquad (11)$$

with K_{ij} being the kernel from the Eq. (2) and C_{ij} the covariance matrix [16] for the different τ values of $\bar{G}(\tau)$. In (9), $\bar{A}(\alpha)$ satisfies

$$\frac{\delta Q}{\delta A} = 0 \qquad (12)$$

where $Q = \alpha S - \chi^2/2$. Details are given elsewhere [16].

The choice (10) of the likelihood function assumes that the data are statistically independent and Gaussian-distributed. These assumptions, which are implicit in a least-squares procedure, are not naturally satisfied. Promoting the consistency of the data with them was achieved by using large bins to reduce the correlations between binned measurements and a large number of bins to generate the Gaussian behavior. The large number of binned measurements also reduces the statistical error associated with the measurements. The number of bins calculated was found empirically to be the number needed so the results did not change if this number was increased.

The Monte Carlo method we used to calculate $G(\tau)$ was originally developed to treat the single-impurity, spin-degenerate (N=2) Anderson model [11] and was later generalised to treat the doubly spin-degenerate Anderson model [18]. It has a number of outstanding advantages. It embeds the impurity in an infinite medium, so there are no finite-size effects, is stable at low temperatures, $\beta \leq 200$, so the simulations can be performed below the Kondo temperature, has almost no sign problem for $N_{imp} \leq 2$, so averages without large variances are obtainable at low temperatures, and produces the Green's function $G(\tau)$ as the natural product. Unfortunately, the operation count scales as $N(N-1)/2$ so we pay quite a price in going from $N = 2$ to $N = 6$. To equilibrate the system and decorrelate the measurements, we unfortunately found that as N is increased, it was necessary to scale up the equilibration and decorrelation times used for $N = 2$ by a factor of N. When combined these scalings mean that simulations for $N = 2$ that took several hours of Cray time would take several days for $N = 6$. The $N = 6$ simulations are beyond what is reasonable on our "central computing facility." Because of this, we turned to parallel computing to beat the scalings.

3. Parallelisation

The details of the Hirsch-Fye algorithm afford several different strategies for parallelisation, all of which, however, require distributing parts of the calculations over all the processors and passing data (messages) on a regular basis

from processor to processor. For some strategies, message passing is frequent and only a small amount of floating point operations are done after the messages are passed. To avoid the communication overhead, we decided to exploit the inherently parallel nature of almost all Monte Carlo calculations. Since our code can run on a single processor, we gave each processor a copy and a separate random number seed, had it run the code, and then collected data from all processors. To see some of the issues involved in this strategy, we will first consider the general aspects of doing a Monte Carlo simulation [19].

In Monte Carlo, we want thermodynamic average such as

$$\langle A \rangle = \frac{\int dx\, A(x)\rho(x)}{Z} \tag{13}$$

where

$$Z = \int dx\, \rho(x) \tag{14}$$

is the partition function (a normalisation constant). The basic property of a Monte Carlo method is to replace the thermal average by a sample average

$$\langle A \rangle \approx \bar{A} \tag{15}$$

where

$$\bar{A} = \frac{1}{M} \sum_{i=1}^{M} A_i \tag{16}$$

and M is the number of measurements A_i of A. If M is large enough and the A_i are *statistically independent* estimates of A, then error estimates are taken to be $\pm m\sigma/\sqrt{M}$ where m is a small integer, usually equal to 1 and

$$\sigma^2 = \frac{1}{M-1} \sum_{i=1}^{M} (A_i - \bar{A})^2 \tag{17}$$

With our parallelisation strategy, the basic Monte Carlo relation (16) becomes

$$\bar{A} = \frac{1}{M}\Big[\underbrace{\sum_{i=1}^{M/n} A_i}_{\text{proc 1}} + \underbrace{\sum_{i=M/n+1}^{2M/n} A_i}_{\text{proc 2}} + \cdots + \underbrace{\sum_{i=(n-1)M/n+1}^{M} A_i}_{\text{proc n}}\Big] \tag{18}$$

however, proper error estimation precludes collecting from each processor only the sum of its measurements. To be able to test for correlations between successive measurements, one needs to collect the measurements from a given processor sequentially and combine these measurements with those from the

other processors into a single file (or buffer). Then, after all node processes are finished, proper error estimates are attempted. If successive measurements from a given processor are correlated and the measurements from the processors are interspersed, then in the combined set of measurements the distance between decorrelated measurements is increased. By grouping measurements by processors, one can test if successive measurements are correlated, and if they are, then one can easily "rebin" successive pairs, triplets, etc. to produce statistical independence.

All the calculations reported here were done on the LANL/IBM workstation cluster which has 16 IBM RISC/6000-560 workstations, each with 128 MB memory, connected ethernet and FDDI networks. For our message passing software, we used PVM [20] because it has the functionality we needed and it is free and easy to use.

Because of cluster usage, the time each processor takes to do M/n measurements can vary widely. To *load balance*, we start the simulation on each node in an infinite loop and whenever a node completes a measurement, the measurement is sent to the host process which writes the result sequentially as a record in a direct-access file and records in a index array which processor wrote to that record. When the host records the desired number of measurements, it terminates all node processes and then uses the index to rewrite the data into another file in proper processor sequence. The number of measurements made varies from node to node but each node does the best it can under the constraints of the time sharing system.

With this overall strategy virtually all inter-processor communication is eliminated. Our code runs 3/2 to 3 times faster on the cluster than on one processor of a Cray-YMP computer, depending on whether the inner loop vector lengths are long (low temperatures) or short (high temperatures). For long loop lengths, the code runs at 250 Mflops sustained on one processor of Cray Y-MP computer. On the cluster, the code has run as long as three days.

4. Results

We calculated the magnetic susceptibility and the single-particle spectral density of the degenerate, single-impurity Anderson model for degeneracies $N = 2, 4$ and 6 in the particle-hole symmetric case and for the single degeneracy $N = 4$ in the asymmetric case [10]. Our results in general agree well with those obtained by the NCA and NRG methods.

As shown if Fig. 1, in the particle-hole symmetric case the magnetic susceptibilities follow the universal Kondo susceptibility curve [3] at low temperatures. The local moment in the intermediate temperature regime is suppressed in comparison to the $N = 2$ case, which is mainly a consequence of the higher degeneracy. The Kondo temperatures, extrapolated from the susceptibility

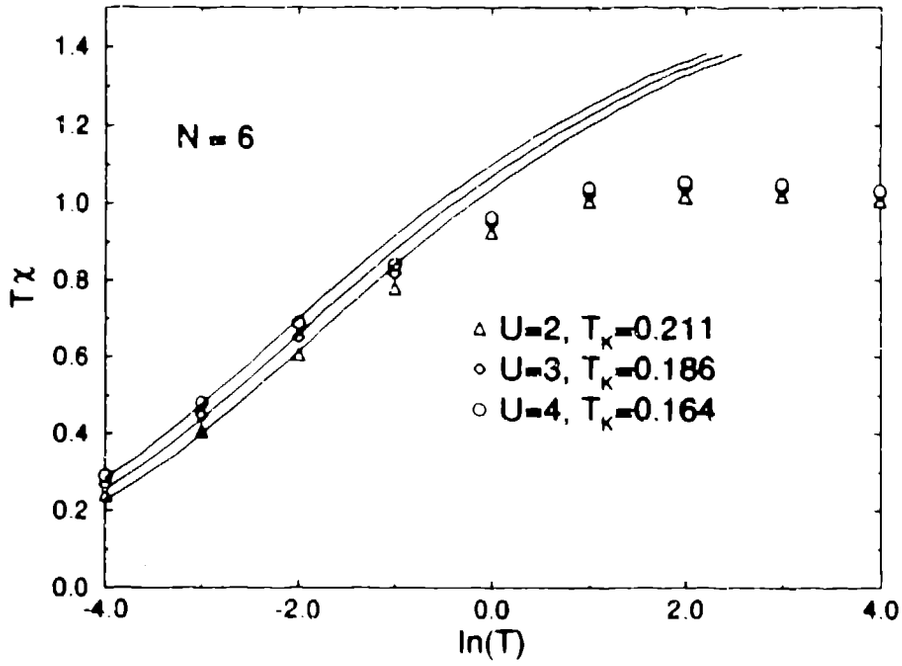

Figure 1. The susceptibility $T\chi$ for $N = 6$ as a function of $\ln(T)$ in units of its high temperature value 3/8. The Monte Carlo results $U = 2$, 3, and 4 are marked by a $\triangle$, $\diamond$, and $\circ$. The solid lines are the universal curves from [3]. T_K is the Kondo temperature. The estimated statistical errors are smaller than the height of the markers.

curves, follow within 10% the estimates from the generalised semi-empirical equation based on the high-temperature perturbational theory of Haldane [21]. In the $N = 4$ case we found good agreement with T_K calculated by Lin and Hirsch [18].

The overall behavior of the single-particle spectral functions follows the predictions of NCA and NRG. The height of the central Kondo peak increases with increasing of degeneracy since the Kondo temperature also increases with degeneracy when the other parameters remain constant. The height of the side peaks scales with degeneracy as $1/N$ as predicted by Zhang and Lee [22]. The position of the side peaks depends slightly upon degeneracy when the unrenormalised energy difference between underlying states remains constant. We associate this degeneracy-induced increased renormalisation with the increase of the effective hybridisation. Within the computational error, we found universal behavior in the spectral functions at low frequencies when $N = 4$. For $N = 2$, such universality had been previously shown by Silver *et al.* [13].

In the asymmetric case, the magnetic susceptibility also follows the uni-

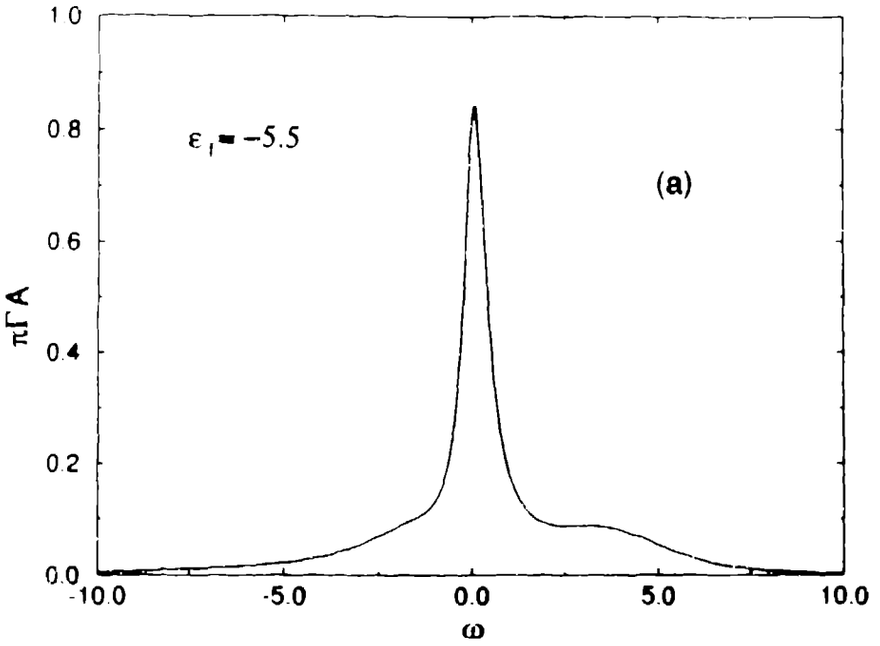

Figure 2. The spectral density $\pi\Gamma A$ in the asymmetric case as functions of the frequency ω calculated for $N = 4$, $U = 4$, $\Gamma = 0.5$, $T = 0.125$, and $\epsilon_f = -5.5$. The symmetric case would correspond to $\epsilon_f = -6$.

versal curve at low temperatures. Because of the highe degeneracy and strong hybridisation, we did not obeerve a pronounced local moment or mixed valence regime. The position of the Kondo peak in the spectral density moves towards positive frequencies as the impurity energy increases from its value at the symmetric case. This shift scales with the Kondo temperature and with the average impurity occupancy as well. The side peaks also shift as the impurity energy increases. Moreover, as the lowest energy state on the impurity site changes from f^2 to f^1 the peak on the BIS side ($\omega > 0$) grows while the peak on the PES side ($\omega < 0$) diminishes because of different phase spaces for transitions $f^1 \rightarrow f^0$ and $f^1 \rightarrow f^2$. An example of an asymmetric spectral density is shown in Fig. 2.

We are just beginning to investigate the effect of the crystal-field and spin-orbit splittings on the spectral functions for the case of $N = 6$. Special focus is on the effect of the splittings on the Kondo peak.

5. Concluding Remarks

By using parallel computing, in an embarassingly simple but effective way, we performed a series of simulations of the degenerate, single-impurity Anderson model [10]. Then, by using methods of Bayesian statistical inference, coupled with the principle of maximum entropy, we extracted from the imaginary-time Green's function the single-particle spectral densities of the model. With these spectral densities, we established the qualitative consistency of NCA calculations. We studied the behavior of the spectral densities as a function of temperature and degeneracy for a variety of model parameters. With $N = 6$ we have a situation that begins to model the actual degeneracy of a Ce impurity. We have begun to mimic crystal effects by splitting the degeneracy of the impurity level.

While our parallelisation scheme exploits one feature common to Monte Carlo simulations, such simulations have other common features: an initialisation, equilibration, and a measurement phase. What we have done is to reduce the computation time by farming out the measurement phase to a large number of processors. Normally, the computing time for the measurement phase dominates the combined computing time for the other two phases. With the current parallelisation strategy, each processor initialises and equilibrates and hence these phases are fixed overheads to the simulations. As the number of processor increases, a point is reached where these two phases dominate the measurement one so that increasing the number of processors decreases the computation time only slightly. Improving performance needs to address the parallelisations of these phases if a large number of processors are to be used. We are currently moving over simulations to the CM-5 computer where we can expect to have 64 to 256 processors for production use.

For the impurity algorithm, the initialisation phase performs a sum like $\sum_n F(n)$. Hence, it is easily parallelised by a procedure (18) similar to the one used for the measurement phase. The computational core of the equlibriation (and measurement) phase is the updating of matrix by an outer product of two vectors. How best to do this is more delicate. Parallelising the equibriation phase appears to have at least two options:

1. Perform the initialisation and equilibriation on a faster computer and feed the results to the cluster or CM-5. A Cray-YMP should perform the initialisation and equilibration by at least a factor of 10 faster than one RISC/6000-560.
2. Distribute the calculation over all the processors in coarse-grained pieces. Instead of each processor providing a separate initialisation, the processors combine to provide just one. While this calculation may have significant interprocessor communication, the use of many processors will be faster than just using one.

With either of these options, all processors then perform separate Monte Carlo

runs, but now after a much shorter equilibration phase and with separate random number seeds, they while become independent processes.

With the CM-5, the vectors units on each processor have just become accessible in the message passing mode. How to exploit them is also being studied. It appears that non-assembley lanuage use of these units will require one's computer code be written in the parallel languages CMFortran or C*. Faced with code rewriting, a natural question to ask, "Can a programming model be implemented in CMFortran or C* that is just as effective but avoids the message passing software althogether?" We are also currently investigating this possibility. Basically, our thinking is still evolving on how to further parallelise our code.

Acknowledgments

We thank S. Hodson and K. Heins-Winkler for their assistence in using the cluster of workstations. This workstation cluster is a joint Los Alamos and IBM project. Our work was supported by the U.S.Department of Energy.

References

1. P.W. Anderson, Phys. Rev. **124**, 41 (1961).
2. A.M. Tsvelik and P.B. Wiegmann, Adv. Phys. **32**, 453 (1983).
3. H.R. Krishna-murty, J.W. Wilkins, and K.G. Wilson, Phys. Rev. B **21**, 1003 (1980); 1044 (1980).
4. L.N. Oliveira and J.W. Wilkins, Phys. Rev. B **24**, 4863 (1981).
5. H.O. Frota and L.N. Oliveira, Phys. Rev. B **33**, 7871 (1986).
6. O. Sakai, Y. Shimisu, and T. Kasuya, J. Phys.Soc. Jpn. **58**, 3666 (1989).
7. N.E. Bickers, D.L. Cox, and J.W. Wilkins, Phys. Rev. B **222:36**, 2036 (1987); N.E. Bickers, Rev. Mod. Phys. **59**, 845 (1987). For alternative approaches to the large-degeneracy expansion see for example, O. Gunnarsson and K. Schönhammer, Phys. Rev. B **28**, 4315 (1983); N. Read and D.M. Newns, J. Phys. C. **16**, 3273 (1982).
8. J.W. Allen, Adv. Phys. **35**, 275 (1986).
9. J.J. Joyce *et al.*, Phys. Rev. Lett. **68**, 236 (1992).
10. J. Bonča and J.E. Gubernatis, Phys. Rev. B, to appear.
11. J.E. Hirsch and R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986).
12. G.D. Mahan, *Many-Particle Physics* (Plenum, New York, 1983), chapter 3.
13. R.N. Silver, D.S. Sivia, J.E. Gubernatis, and M. Jarrell, Phys. Rev. Lett. **65**, 496 (1990)
14. A. Papoulis, *Probability and Statistics* (Prentice-Hall, Englewood Cliffs,

1990).
15. This identification follows from functionally integrating Bayes theorem with respect to A, using $\int \mathcal{D}A \Pr[A|\tilde{G}] = 1$, and then observing that $\Pr[\tilde{G}] = \int \mathcal{D}A \Pr[\tilde{G}|A] \Pr[A]$.
16. J.E. Gubernatis, R.N. Silver, and M. Jarrell, Phys. Rev. B **44**, 6011 (1991).
17. R.K. Bryan, Eur. Biophys. J. **18**, 165 (1990).
18. H.Q. Lin and J.E. Hirsch, Phys. Rev. B **37**, 1864 (1988).
19. M.H. Kalos and P.A. Whitlock, *Monte Carlo Methods* (Wiley, New York, 1986).
20. A. Geist, et al., "PVM User's Guide, ver. 2.3," Oak Ridge National Laboratory report ORNL/TM-12187.
21. S.D.M. Haldane, J. Phys. C **11**, 5015 (1978)
22. F.C. Zhang and T.K. Lee, Phys. Rev. B **30**, 1556 (1984).