LA-UR-91- 3290

LA-UR--91-3290

5--1

DE92 002376

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

NOV 06 1991

TITLE: BAYESIAN METHODS, MAXIMUM ENTROPY, AND QUANTUM
MONTE CARLO

AUTHOR(S): J. E. Gubernatis
R. N. Silver
Mark Jarrell*

SUBMITTED TO: Proceedings: "Computation Physics for Condensed Matter Phenomena,"
October 21-23, 1991, Osaka, Japan.
To be published by Plenum Press

*University of Cincinnati, Cincinnati, OH 45221

By acceptance of this article, the publisher recognizes 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.

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2829 5/81

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

782- MASTER

Bayesian Methods, Maximum Entropy, and Quantum Monte Carlo

J. E. Gubernatis, Theoretical Division
Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.

R. N. Silver, Theoretical Division and Los Alamos Neutron Scattering Center
Los Alamos National Laboratory, Los Alamos, NM 87545, U.S.A.

Mark Jarrell, Department of Physics
University of Cincinnati, Cincinnati, OH 45221, U.S.A.

Abstract. We heuristically discuss the application of the method of maximum entropy to the extraction of dynamical information from imaginary-time, quantum Monte Carlo data. The discussion emphasizes the utility of a Bayesian approach to statistical inference and the importance of statistically well-characterized data.

1. Introduction

Recently, we proposed procedures that for the first time permit the consistent extraction of dynamical properties from imaginary-time quantum Monte Carlo data. The technical details of these procedures, as well their application to several different physical models, studied by several different quantum Monte Carlo methods, have been published [1,2]. In these proceedings, we will summarize our current understanding of how these procedures work by discussing them in a more heuristic manner than previously possible.

Most quantum Monte Carlo simulations at finite temperatures are done in imaginary time. The Wick rotation, $it \rightarrow \tau$, converting real time t to imaginary time τ , transforms oscillatory exponentials e^{-itH} of the Hamiltonian H into diffusive ones, $e^{-\tau H}$. Monte Carlo sampling of a function that becomes highly oscillatory at large, real times is inefficient as sampling on smaller and smaller times scales becomes necessary just to achieve proper self-cancellations [3]. In most cases, however, it is a spectral density, which describes the distribution of elementary excitations associated with various one- and two-particle Green's functions, and not the real-time behavior which is generally sought. Still again, the difficulty in obtaining highly oscillatory, long-time information to insure proper low-frequency behavior is inhibiting.

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.

In imaginary-time, a different difficulty arises. With quantum Monte Carlo estimates $\tilde{G}(\tau)$ of a one- or two-particle Green's function $G(\tau)$, the generic problem is to find the associated spectral density $A(\omega)$ from

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

where the $\pm$ refers to Fermi and Bose statistics and $\tau > 0$. The difficulties arise at large positive and negative frequencies because of the exponential behavior of the kernel

$$K(\tau, \omega) = \frac{e^{-\tau\omega}}{1 \pm e^{-\beta\omega}}. \quad (2)$$

For the forward problem, i.e., "Given A , what is G ?", $G(\tau)$ is insensitive the large $|\omega|$ details of $A(\omega)$. For the inverse problem, i.e., "Given G , what is A ?", this insensitivity leads to an ill-posed problem, so with incomplete and inaccurate information about $G(\tau)$, an infinite number of solutions exist. The task is then to select from these solutions one that is best by some criterion.

Previously, several criteria have been proposed. None lead to a satisfactory procedure with the potential of handling the more difficult problems. The procedures we have proposed [1] succeed because they use more relevant prior information about the behavior of $A(\omega)$ and are accompanied by a closer assessment of the statistical character of the data produced by the quantum Monte Carlo method. Their structure is completely tied to the Bayesian approach to statistical data analysis and the use of the principle of maximum entropy.

2. Bayesian Inference and Maximum Entropy

Given two events a and b , Bayes's theorem says [4]

$$\Pr[a, b] = \Pr[a|b] \Pr[b] = \Pr[b|a] \Pr[a] \quad (3)$$

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 . Also from probability theory, we have the condition for the marginalization of a variable

$$\Pr[a] = \int db \Pr[a, b]. \quad (4)$$

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 $\tilde{A}(\omega)$ that maximizes $\Pr[A|\tilde{G}]$. We take

$$\Pr[A|\tilde{G}] = \Pr[\tilde{G}|A] \Pr[A] / \Pr[\tilde{G}]. \quad (5)$$

$\Pr[A|\tilde{G}]$ is called the posterior probability, $\Pr[\tilde{G}|A]$, the likelihood function, and $\Pr[A]$, the prior probability. $\Pr[\tilde{G}]$ is a normalization constant [5], called the evidence.

Since the spectral density $A(\omega)$ is non-negative and normalizable to unity, it can be interpreted as still another probability function. In the absence of data, the principle of maximum entropy [4] says that the values of a probability function $A(\omega)$ are to be assigned by maximizing the entropy expression

$$S = - \int d\omega A(\omega) \ln A(\omega).$$

In the absence of data (or with the lack of information about the likelihood function), $\Pr[A|\tilde{G}] \propto \Pr[A]$. Hence, the A that maximizes the posterior probability is the same as the one that maximizes S when

$$\Pr[A] \propto e^{\alpha S}. \quad (6)$$

Thus, the entropy associated with the probability function fixes the prior probability up some constant α [6].

If $A(\omega)$ is not normalized, we could maximize

$$S - \lambda \left(1 - \int d\omega A(\omega) \right)$$

where λ is a Lagrange multiplier, or equivalently, we can assign probabilities, i.e., find $A(\omega)$, by maximizing

$$S = - \int d\omega A(\omega) \left(\ln A(\omega) - A(\omega) \right).$$

In the literature [6, 7], a fuller form for S is used

$$S \equiv - \int d\omega A(\omega) \left(\ln[A(\omega)/m(\omega)] - A(\omega) + m(\omega) \right). \quad (7)$$

The function $m(\omega)$ is called the default model, because in the absence of data, S is maximized when $A(\omega) = m(\omega)$. From one point of view, $m(\omega)$ sets the zero of the entropy. The importance and use of the default model, however, goes beyond this, as we will shortly discuss.

In the presence of data, i.e., the values of $\tilde{G}(\tau)$ for N different values of $\tau = \tau_i$, we could assign probabilities by maximizing

$$S - \sum_{i=1}^N \lambda_i \left(\tilde{G}(\tau_i) - \int d\omega K(\tau_i, \omega) A(\omega) \right).$$

This procedure ignores any errors associated with the data, but these errors are important information needed to determine confidence in the results. The errors associated with the data are best incorporated into the analysis through the likelihood function.

2. Some Intermediate Observations

To motivate about the default model a bit more, we will assume that we know $A(\omega)$. Interpreting $A(\omega)$ as a probability function, we could fix τ in (1), evaluate the integral by a Monte Carlo procedure where we generate a sequence of ω_i drawn from the probability distribution $A(\omega)$, estimate $G(\tau)$ by

$$\bar{G}(\tau) = \frac{1}{M} \sum_{i=1}^M K(\tau, \omega_i),$$

and then assign an error σ to this estimate from

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M \left(K(\tau, \omega_i) - \bar{G}(\tau) \right)^2. \quad (8)$$

In Monte Carlo integration, importance sampling [8] is often used to reduce variance by introducing a trial function $m(\omega)$ that is similar to the actual function, rewriting the integral as

$$G(\tau) = \int_{-\infty}^{\infty} d\omega m(\omega) K(\tau, \omega) \frac{A(\omega)}{m(\omega)},$$

and drawing the random variables ω_i from $m(\omega)$. The estimate of $\bar{G}(\tau)$ becomes

$$\bar{G}(\tau) = \frac{1}{M} \sum_{i=1}^M K(\tau, \omega_i) \frac{A(\omega_i)}{m(\omega_i)}.$$

If $m(\omega) = K(\tau, \omega)A(\omega)/\bar{G}(\tau)$, it is easy to show that for large M the correct value of $\bar{G}(\tau)$ is produced with zero variance. Of course, this method of eliminating variance requires *a priori* knowledge of the exact answer, but clearly by using relevant, prior information about the expected answer, we can produce a more certain result. On the other hand, an unwise choice for $m(\omega)$ can be counterproductive.

Our operational point-of-view is that the default model is a "variance reduction" tool analogous to the trial function in Monte Carlo integration: the closer the default model is to answer, the better will the answer be. We are not trying to assume the answer, but to solve an ill-posed problem, we need to use as much the prior information as possible. The default model is a tool towards this end. In many cases, we may not know much before hand.

It is clear that the Monte Carlo integration scheme just described produces independent estimates of $G(\tau)$ at different τ values. At a given τ , successive values of ω_i , and hence of the estimators $K(\tau, \omega_i)$ of $G(\tau)$, can be expected to

be reasonably statistically independent. The error estimate (8) assumes that these independent values are Gaussian distributed, but clearly their distribution depends on $K(\tau, \omega)$ and $A(\omega)$. This estimate σ is meaningful only if M is sufficiently large so that the central limit theorem applies [4].

In the quantum Monte Carlo applications, we are addressing the inverse problem, not the forward problem just described. As we will discuss, for different values of τ , the $\tilde{G}(\tau)$ are not independent so care must be taken to achieve statistically independent estimates data. Reliable error measurements will also depend on the central limit theorem, but one has to face the fact that this theorem does not specify the practical size of M needed to know when the theorem applies. The need to consider this question is dictated by the choice of the likelihood function used. Almost always, a specific choice is made.

3. Likelihood Function

The method of maximum likelihood [9] is commonly used in parameter fitting problems. For simple probability functions, like a Gaussian, fitting two parameters to the data, the mean and standard deviation, is sufficient to determine completely the probability function. Using the data to find $A(\omega)$ is type of parameter fitting problem in as much as we will determine a set $A = (A_1, A_2, \dots, A_N)$ of values of A at a number of discrete values ω_i of ω , but the number N of parameters we determine is on the order of 200. From (1), we find that a given A makes a specific prediction for a set of $G(\tau)$ values $G = (G_1, G_2, \dots, G_L)$ at various discrete values τ_i of τ . In the method of maximum likelihood the objective is to vary A so that the resulting G maximizes some likelihood function.

In maximum likelihood method, the $\tilde{G}(\tau)$ are regarded as random variables drawn from a probability distribution $f(\tilde{G}; G)$ specified by G . For a single random sampling of $\tilde{G}(\tau)$, a set of random variables $\tilde{G}^{(j)} = (\tilde{G}_1^{(j)}, \tilde{G}_2^{(j)}, \dots, \tilde{G}_L^{(j)})$ describing the values of $\tilde{G}$ at discrete values τ_i of τ is realized. For a single $\tilde{G}^{(j)}$, we can assign a number $dP^{(j)} = f(\tilde{G}^{(j)}; G)dG$. For M independent measurements, $dP = \prod_{j=1}^M f(\tilde{G}^{(j)}; G)dG$. The *logarithm* of the likelihood function is $L = \sum_{j=1}^M \ln f(\tilde{G}^{(j)}; G)$. The method of maximum likelihood is based on the assumption that a single peak in the space of the parameters G dominates this function so that the mode $\tilde{G}$ is a reliable estimate of the parameters to be fitted and the width of the peak provides some measure of confidence in the fit.

From a Bayesian point of view, the method of maximum likelihood takes $\Pr[A|\tilde{G}] \propto \Pr[\tilde{G}|A]$ so that the prior probability $\Pr[A]$ is ignored. Hence, finding the A that maximizes the posterior probability is equivalent to finding the A that maximizes $\Pr[A|\tilde{G}] \propto e^L$. The method does not specify the function $f(\tilde{G}^{(j)}; G)$. In some applications, it might be known, but it is definitely an un-

known in the quantum Monte Carlo case. When the number of measurements becomes large, by central limit theorem arguments, modes becomes means, etc., and the asymptotic behavior becomes known [9]

$$e^L \propto e^{-\frac{1}{2}\chi^2} \quad (9)$$

where

$$\chi^2 = \sum_{i,j}^L (\bar{G}_i - G_i) [C^{-1}]_{ij} (\bar{G}_j - G_j) \quad (10a)$$

and

$$\bar{G}_i = \frac{1}{M} \sum_{j=1}^M \bar{G}_i^{(j)}. \quad (10b)$$

C is the covariance matrix

$$C_{ik} = \frac{1}{M(M-1)} \sum_{j=1}^M (\bar{G}_i - \bar{G}_i^{(j)}) (\bar{G}_k - \bar{G}_k^{(j)}). \quad (10c)$$

Maximizing the likelihood is equivalent to minimizing χ^2 . A least-squares fitting procedure (10) is thus a special case of the method of maximum likelihood with $L = -\frac{1}{2}\chi^2$.

The least-squares method has several conditions for its proper application: independent measurements and a multi-variant Gaussian distribution of the data. It also makes an important assumption that the prior probability is unimportant or unknown. What is curious about least-squares determinations of the spectral density is that one of the first things done, when undesirable solutions were obtained, was to seek new solutions by constraining the fit to be positive and normalized. These conditions, however, are inherent to the entropic prior (6). The more significant constraint imposed, however, was the one of smoothness, which is *ad hoc* because there is no *a priori* reason to believe that the spectral density has such a property. When combined with the likelihood function, the entropic prior has a smoothing effect on the result that is principally dictated by the data and not by the user of the data.

4. Bayesian Inference Revisited

With (6) and (9), we express Bayes's theorem (5) as $\Pr[A|\tilde{G}] \propto e^Q$ where

$$Q = \alpha S - \frac{1}{2}\chi^2. \quad (11)$$

Finding the $A(\omega)$ that maximizes $\Pr[A|\tilde{G}]$ is thus equivalent to maximizing Q with respect to $A(\omega)$

$$\left. \frac{\delta Q}{\delta A} \right|_{A=\hat{A}} = 0 \quad (12)$$

The result, however, will depend on α , and in what is called *historic maximum entropy* [1, 6, 7], α is adjusted to make $\chi^2 = N$. This choice is the expected value of χ^2 when the errors are Gaussian noise but is otherwise *ad hoc*.

We now describe a recent development, called *classic maximum entropy* [6, 7]. We start by rewriting (5) as

$$\begin{aligned} \Pr[A, \alpha|\tilde{G}] &= \Pr[\tilde{G}|A, \alpha] \Pr[A, \alpha] / \Pr[\tilde{G}] \\ &= \Pr[\tilde{G}|A, \alpha] \Pr[A|\alpha] \Pr[\alpha] / \Pr[\tilde{G}] \end{aligned} \quad (13)$$

to make explicit the dependence of our results on α . The normalization factor $\Pr[\tilde{G}]$ equals $Z_L Z_S$ where the factors of Z normalize the likelihood function and the entropic prior. Z_S depends on α [7]. Introduced is a new function $\Pr[\alpha]$, which we will usually take to be what is called the Jeffrey prior [1]. In (13), the first two functions on the right-hand side are just (9) and (6).

Using (4) and then (3), we can also write the posterior probability $\Pr[A|\tilde{G}]$ as

$$\begin{aligned} \Pr[A|\tilde{G}] &= \int d\alpha \Pr[A, \alpha|\tilde{G}] \\ &= \int d\alpha \Pr[A|\tilde{G}, \alpha] \Pr[\alpha|\tilde{G}]. \end{aligned} \quad (14)$$

The distribution function $\Pr[\alpha|\tilde{G}]$ follows from functionally integrating (13) with respect to $A(\omega)$

$$\begin{aligned} \Pr[\alpha|\tilde{G}] &= \int \mathcal{D}A \Pr[\tilde{G}|A, \alpha] \Pr[A|\alpha] \Pr[\alpha] / \Pr[\tilde{G}] \\ &= \Pr[\alpha] \int \mathcal{D}A A(\omega) \frac{e^{-Q}}{Z_L Z_S}. \end{aligned} \quad (15)$$

If the number of data is large, then $\Pr[\alpha|\tilde{G}]$ is often sharply peaked at some value of $\alpha = \hat{\alpha}$ [7], and (14) becomes $\Pr[A|\tilde{G}] \propto \Pr[A|\tilde{G}, \hat{\alpha}]$. With the $\hat{\alpha}$, the $A = \hat{A}$ satisfying (12) becomes the classic maximum entropy solution. The ill-posed nature of the our problem, however, restricts the space in which we can search for a solution so that our effective number of data is usually small and we cannot legitimately use the above.

To proceed, for a fixed value of α , we find the solution $\hat{A}_\alpha$ that satisfies (12). Then, we compute numerically the solution $\bar{A}(\omega)$ defined by

$$\bar{A}(\omega) \equiv \int d\alpha A(\omega) \Pr[A|\alpha] \quad (16)$$

where $\text{Pr}[\alpha|\bar{G}]$ is given by (15). In the absence of data, this procedure produces the default model as the answer.

These are the major details of classic maximum entropy method and its application to quantum Monte Carlo data; the seminal works [6, 7, 10] should be discussed for a more detailed description. The main point is that α is placed on a similar footing as $A(\omega)$. This allows $A(\omega)$ to be determined without needing to adjust α or using properties of α not supported by the data.

5. Quantum Monte Carlo

The procedures just discussed are not tied to the problem of extracting spectral densities from imaginary-time quantum Monte Carlo data. They are quite general and can be applied to the solution of almost any ill-posed problem where the function sought is non-negative. To tie them to our problem, we must make certain that our measurements $\bar{G}^{(j)}$ are statistically independent and distributed by a multivariate Gaussian probability function. These conditions have to be promoted because they are not even roughly mimicked by the quantum Monte Carlo data [1].

Normally, the simulation produces a sequence of measurements that are partitioned into M groups (bins), each of length k . The distribution from which these measurements are drawn is not known. The idea is create a distribution that is known by making k large enough so that the means of the measurements in each bin, i.e., the $\bar{G}^{(j)}$, are statistically independent and by making M large enough so that these means are approximately distributed according to the Gaussian distribution predicted by the central limit theorem.

Auto-correlation measures exist that are helpful in estimating the distance between correlated measurements [11], but more than one correlation length is most likely present, so such estimates are rough and the actual bin size k effective for decorrelation needs to be 10 to 20 times larger than the estimated correlation length. Additionally, the behavior at $\tau = 0$ generally differs from the behavior at $\tau = \frac{1}{2}\beta$, so this measure, as the ones to be discussed below, needs to be done at several different τ values and the worst case determines the bin-size to be used.

With M independent measurements, their mean, their variance, and an histogram of the measurements, a standard χ^2 test can be applied to test the consistency of the histogram with a normal distribution with the same mean and variance [1]. Such a test is helpful but is more useful for establishing inconsistency with a normal distribution than consistency. The lowest order corrections to the normal distribution predicted by central limit theorem are controlled by the skewness and kurtosis of the distribution of the measurements [12]. If this distribution is Gaussian, then these deviations would be zero, but in a statistical sampling process, these measures, as the standard deviation,

are expected to scale as $1/\sqrt{M}$ [13]. It is important to increase M until this scaling is approximately obeyed [14].

There are two matrices in the problem: the $N \times N$ covariance matrix C and the $N \times L$ discretized form of the kernel K . Both matrices can have ratios of singular values that exceed what can be handled accurately on a finite precision computer. When the singular values become too small, the rank of these matrices has to be reduced to maintain numerical stability. These reductions represent a loss of information. For C , the reduction says that not all the data represent independent information. For K , it is a reflection of the ill-posed nature of the problem. For C , the reduction is greatly minimized if $N > M$ [14]. For K , it is generally severe, and K becomes a matrix of rank s where $s \ll L, N$.

The covariance matrix is poorly approximated by a diagonal matrix. In the illustrative example of Monte Carlo integration, the estimates of $G(\tau)$ at different values of τ were produced independently which is not the same as saying they are independent. The independent production is a consequence of the algorithm. In quantum Monte Carlo algorithms, the Green's function at one value of τ is generally produced from all the others.

6. Concluding Remarks

We have tried to present a discussion bridging the seminal works in Bayesian methods and maximum entropy [6, 7, 10] with their application to the extraction of dynamical information from imaginary-time quantum Monte Carlo data [1, 2]. These methods provide a consistent logical framework to approach the problem with the assumptions in the approach clearly defined. The Bayesian methods lead to a procedure (16) with no adjustable parameters.

Ill-posedness causes the need for such sophisticated procedures. By the nature of ill-posedness, the solution of an ill-posed problem is sensitive to the size of the errors on the input data. The procedures described require not only small errors but also Gaussian-distributed ones. The need for Gaussian behavior is a result of a particular choice of the likelihood function. Other choices could have been made, but they would be inconsistent with the way quantum Monte Carlo data, or data from other simulation methods, are usually analyzed. Since producing data with the proper statistical qualities is the hardest part of the problem, investigating other likelihood functions is certainly warranted. Because a specific function is unlikely to follow from analytic considerations, such a function will most likely need to be generated empirically from the quantum Monte Carlo data itself by use of non-parametric, robust statistical methods.

References

1. J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, "Quantum Monte Carlo simulations: dynamics from imaginary-time data," *Phys. Rev. B*, to appear.
2. R. N. Silver, D. S. Sivia, and J. E. Gubernatis, *Phys. Rev. B* **41**, 2380 (1990); R. N. Silver, J. E. Gubernatis, D. S. Sivia, and M. Jarrell, *Phys. Rev. Lett.* **65**, 496 (1990); M. Jarrell, J. E. Gubernatis, R. N. Silver, and D. S. Sivia, *Phys. Rev. B* **43**, 1206 (1991); M. Jarrell, J. E. Gubernatis, and R. N. Silver, *Phys. Rev. B* **44**, 5347 (1991); J. Deisz, M. Jarrell, D. L. Cox, *Phys. Rev. B* **42**, 4869 (1990); M. Jarrell, D. S. Sivia, B. Patton, *Phys. Rev. B* **42**, 4804 (1990); J. Deisz, K.-H. Luc, M. Jarrell, D. L. Cox, "Spin and charge dynamics for the one-dimensional t-J model," *Phys. Rev. B*, to appear.
3. J. D. Doll, D. L. Freeman, and T. L. Beck, *Adv. Chem. Phys.* **77**, 61 (1990), and references therein.
4. A. Papoulis, *Probability and Statistics* (Prentice-Hall, Englewood Cliffs, 1990).
5. 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 $\Pr[\tilde{G}] = \int \mathcal{D}A \Pr[\tilde{G}|A] \Pr[A]$.
6. J. Skilling, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer, Dordrecht, 1989), pg. 45.
7. S. F. Gull, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer, Dordrecht, 1989), pg. 53.
8. M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods, Volume 1: Basics* (Wiley, New York, 1986), chap. 4.
9. S. Brandt, *Statistical Computational Methods in Data Analysis* (North-Holland, New York, 1983).
10. R. K. Bryan, *Eur. Biophys. J.* **18**, 165 (1990).
11. G. Jacucci and A. Rahman, *Nuovo Cimento D* **4**, 341 (1983).
12. D. T. Gillespie, *Am. J. Phys.* **51**, 520 (1983).
13. W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986), chap. 13.
14. M. Jarrell, J. E. Gubernatis, and R. N. Silver, "Dynamical properties of the asymmetric, single-impurity Anderson model," in preparation.

END

**DATE
FILMED**

12/19/91