Maximum-entropy method for analytic continuation of quantum Monte Carlo data

R. N. Silver,* D. S. Sivia,* and J. E. Gubernatis[†]
Los Alamos National Laboratory, Los Alamos, New Mexico 87545
(Received 18 September 1989)

An outstanding problem in the simulation of condensed-matter phenomena is how to obtain dynamical information. We consider the numerical analytic continuation of imaginary-time quantum Monte Carlo data to obtain real-frequency spectral functions. This is an extremely ill-posed problem similar to the inversion of a Laplace transform. We suggest an image-reconstruction approach, which has been widely applied to data analysis in experimental research. Specifically, we apply the maximum-entropy method (ME) to the analytic continuation of quantum Monte Carlo data. We report encouraging preliminary results for the Fano-Anderson model of an impurity state in a continuum. The incorporation of additional prior information, such as sum rules and asymptotic behavior, can be expected to significantly improve results. We compare (ME) to alternative methods. We also discuss statistical error propagation for the analytic continuation problem via the likelihood function, which is independent of the choice of image-reconstruction method. This includes the sensitivity of the data to structure in the spectral function, the optimization of Monte Carlo simulations, and how to incorporate covariance in the statistical errors of the Monte Carlo method.

I. INTRODUCTION

The dynamical properties of strongly correlated many-body systems are of fundamental interest in most areas of condensed-matter research. Quantum Monte Carlo methods[1] for the numerical simulation of such systems provide data on thermodynamic Green's functions in imaginary time. While this is appropriate for the determination of thermodynamic properties, it is difficult to extract the dynamical properties from the data because an analytic continuation from imaginary time to real time is required. The transform which relates the imaginary-time data to the real-frequency spectral function is similar to a Laplace transform. If one has physical reasons to believe in a model for the spectral function, the parameters of the model can be determined by fitting to the data. In the absence of a model, which is generally the case, the problem of inferring the spectral function by inverting such transforms can be extremely ill posed when the data are noisy and incomplete.[2]

We suggest that this analytic continuation is essentially an image-reconstruction problem, which is similar to others in a wide variety of experimental fields including radio astronomy, magnetic resonance imaging, photographic image enhancement, neutron scattering, etc. The *image* sought in the present case is the real-frequency spectral function. Almost any of the successful methods for image reconstruction should work with sufficiently good data. Given the limited data usually available, image-reconstruction methods improve as they incorporate more prior knowledge about the quantity of interest. In choosing an image-reconstruction method for our problem, the criteria to be considered should include: (1) the conceptual foundations of the statistical inference approach; (2) the efficiency of the method in making the most out of the prior knowledge and the data available,

with the caution that the method should tend to put structure in the image only if there is statistically significant evidence for it in the data; and (3) the computational cost and programming effort required.

The maximum-entropy method (ME) (Ref. 3) meets the first two criteria admirably, but it can score poorly on the third. For the analytic continuation problem, however, this weakness is of little consequence because the computational cost of ME is negligible compared to the cost of the quantum Monte Carlo calculations required to produce the data. The programming effort required is also negligible because we use an existing third generation commercial code for ME image processing.[4]

ME, like many successful image-reconstruction techniques, can be understood in terms of Bayesian probability theory,[5] also called the law of conditional probabilities, which provides a logically consistent approach to statistical inference. Bayes' theorem encapsulates the learning process: our state of knowledge about some quantity of interest after an experiment (or simulation) depends on both the relevant data and our prior state of knowledge (or the lack thereof).

For the analytic continuation problem, the most important prior knowledge is that the spectral function is a positive additive probability distribution. Additional prior knowledge may include sum rules, symmetry properties, asymptotic behavior, etc. ME is a special case of Bayes' theorem appropriate to such probability distributions. It enforces the positivity of the spectral function, it tends to put structure in the image only if warranted by the data, and for most applications it has no adjustable parameters. When there are adjustable parameters, these can (in principle) also be estimated using Bayesian methods. Because ME is based on probability theory, it can provide error estimates of the reliability of the features in the spectral function given an adequate char-

acterization of the statistical errors in the data. There is a vast body of experience in using ME on diverse image-reconstruction problems,[6] which provides insight into the images ME will produce for our problem.

Figure 1 shows an example of the use of the ME method for photographic image enhancement. The photograph of a "getaway" car is blurred because of car motion observed with a finite camera shutter speed. The blurring function is evident from the streak produced by the point source above the fender. This is used as prior knowledge in a ME reconstruction. While the number on the license plate cannot be read from the original photo, it is easily read from the ME reconstruction.

The present paper describes the application of ME to the analytic continuation problem for Monte Carlo data. We use the spectral function of the Fano-Anderson model of a bound state in a continuum for illustration.[7] We compare ME to other approaches to the analytic continuation problem in the recent literature by Jarrell and Biham,[8] and by White et al.[9] The problem has also been approached by Schuttler and Scalapino[10] using least-squares fitting, and by Hirsch[11] using Padé approximants.

The propagation of statistical errors in the analytic continuation of Monte Carlo data is embodied in the likelihood function, which is independent of the choice of image-reconstruction method. This enables error estimates of the integrated intensities of features in an image, such as is implemented in modern versions of ME.[4] We analyze the likelihood function for the analytic continuation problem to show that the data are most sensitive to sharp structure in the spectral function for frequencies near the chemical potential, and that at higher and lower frequencies only broad structure can be reliably determined. We discuss a likelihood function analysis of the optimization of Monte Carlo calculations. We suggest how the likelihood function might be modified to account for covariance in the statistical errors of quantum Monte Carlo data.

II. THE ANALYTIC CONTINUATION PROBLEM

Quantum Monte Carlo generates the imaginary time (Matsurbara) Green's function

$$G(\tau) \equiv -\langle T_\tau[\hat{a}(\tau)\hat{a}^\dagger(0)] \rangle , \qquad (1)$$

where τ is imaginary time, T_τ denotes time ordering, $\hat{a}^\dagger$ and $\hat{a}$ are particle creation and annihilation operators, and $\langle \ \rangle$ denotes a grand canonical ensemble average over the states of the many-body system. Knowledge of $G(\tau)$ in the range $0 \leq \tau \leq \beta$, where $\beta \equiv 1/k_B T$, is complete because $G(\tau)$ is periodic for bosons and antiperiodic for fermions. The dynamical properties of interest are given by the density of single-particle states, which is termed the spectral function, $A(\omega)$. This is related to $G(\tau)$ by a transform

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

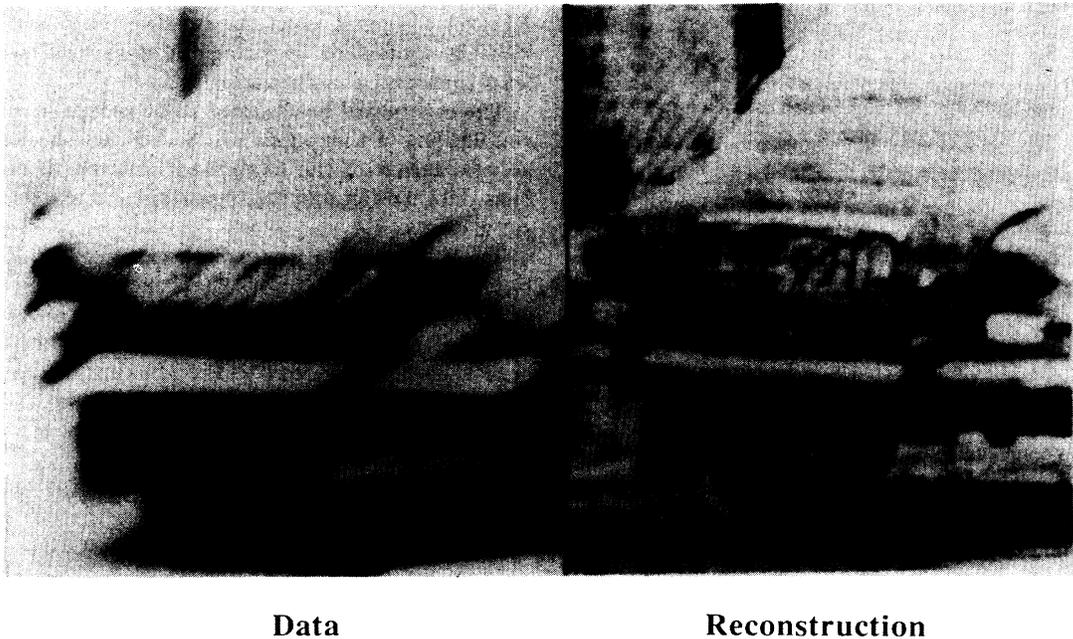Data   Reconstruction

FIG. 1. Optical deconvolution: an example of the use of ME.

where μ is the chemical potential. The problem is to invert this transform in order to determine $A(\omega)$ from $G(\tau)$. Because this involves going from imaginary time to real time, it is equivalent to analytic continuation.

The Monte Carlo calculations provide data, $\{G_d(\tau_k)\}$ (i.e., calculated at a discrete set of τ points, $\{\tau_k\}$, distributed in the range $0 \leq \tau_k \leq \beta$), about $G(\tau)$, which are incomplete and noisy (i.e., subject to statistical error). In this case, the inversion is ill posed because there exist an infinity of $A(\omega)$, all of which fit the data according to a χ^2 measure. Like the inversion of a Laplace transform, the analytic continuation problem is *extremely* ill posed, because very small changes in the data can lead to very large changes in $A(\omega)$. The fitting of an $A(\omega)$ to the data by minimization of χ^2 (the method of least squares), is likely to put statistically insignificant structure into the $A(\omega)$ obtained. This nonuniqueness and extreme sensitivity to error have limited the ability to infer $A(\omega)$ from data produced by quantum Monte Carlo calculations. The data may also be subject to systematic error because of the limitations of the quantum Monte Carlo calculation, but we do not consider these further in this paper.

III. THE MAXIMUM-ENTROPY METHOD

The image-reconstruction approach we adopt is to infer the "best" $A(\omega)$ according to probability theory arguments. The method should tend to put structure in $A(\omega)$ only if there is statistically significant evidence for it in the data. The method must also be capable of providing estimates of the statistical reliability of the results. The method should enforce whatever prior knowledge about $A(\omega)$ we may have, such as positivity and sum rules; i.e.,

$$A(\omega) \geq 0 , \tag{3}$$

and

$$\int_{-\infty}^{\infty} d\omega\, A(\omega) = 1.0 . \tag{4}$$

Bayes' theorem provides our basis for statistical reasoning. It is simply the law of conditional probabilities:

$$P[X,Y] = P[X|Y] \times P[Y] = P[Y|X] \times P[X] . \tag{5}$$

Here, $P[X,Y]$ is the joint probability of X and Y, and $P[X|Y]$ is the conditional probability of X given Y. To specialize this to the analytic continuation problem, three quantities enter into the theorem. The first is the probability distribution of $A(\omega)$ before the experiment (or Monte Carlo calculation) is conducted, $P[A(\omega)]$, which is termed the *prior probability*. For example, since we know that $A(\omega)$ cannot be negative, $P[A(\omega)]$ should be zero for $A(\omega)$ that go negative. The second is the conditional probability of producing the data, $\{G_d(\tau_k)\}$, via the experiment from a given $A(\omega)$, which is termed the

likelihood function, $P[G_d(\tau)|A(\omega)]$. It represents the modification of the prior probability on $A(\omega)$ by the experiment. The third is the conditional probability of $A(\omega)$ given the data after the experiment, $P[A(\omega)|G_d(\tau)]$, which is termed the *posterior* probability. The theorem states that the posterior probability is proportional to the product of the prior probability and the likelihood function, i.e.,

$$P[A(\omega)|G_d(\tau)] \propto P[G_d(\tau)|A(\omega)] \times P[A(\omega)] . \tag{6}$$

The *image* to be presented, $A_I(\omega)$, is the $A(\omega)$ which maximizes the posterior probability. The statistical reliability of the image is to be obtained from the variation of $P[A(\omega)|G_d(\tau)]$ about this maximum.

The likelihood function contains the new information provided by the experiment. If the data, $\{G_d(\tau_k)\}$, are assumed to be independent and Gaussian distributed with error σ_k, the likelihood function is related to the χ^2 measure by

$$P[G_d(\tau)|A(\omega)] \propto \exp(-\chi^2/2) , \tag{7}$$

where

$$\chi^2 = \sum_k \frac{[G_d(\tau_k) - G_f(\tau_k)]^2}{\sigma_k^2} . \tag{8}$$

Here, $G_f(\tau_k)$ are the data that a given choice of $A(\omega)$ would produce in the absence of noise. For example, if the $A(\omega)$ is to be determined in pixels at discrete frequencies, $\{\omega_i\}$, then

$$G_f(\tau_k) \simeq \sum_f A(\omega_i)\Delta\omega \frac{e^{-(\omega_i - \mu)\tau_k}}{1 + e^{-(\omega_i - \mu)\beta}} . \tag{9}$$

If one ignores the prior probability in Eq. (6), i.e., $P[A(\omega)] \rightarrow$ const, then maximizing the posterior probability is equivalent to fitting the data by the method of least squares, i.e., minimizing χ^2.

The errors will be assumed to be independent in all the simulations of this paper. In Sec. V, we shall discuss the generalization of the likelihood function to nonindependent data, which may be important for real Monte Carlo simulations.

The maximum-entropy method corresponds to a particular choice of prior that incorporates the prior knowledge that the spectral function is positive and additive. In that case, a variety of different statistical inference arguments[3,12] lead to the conclusion that the prior should take the form

$$P[A(\omega)] \propto \exp(aS) . \tag{10}$$

Here, S is the information theory entropy[13] of the image, i.e.,

$$S = \sum_i \Delta\omega \left[A(\omega_i) - m(\omega_i) - A(\omega_i)\ln\left[\frac{A(\omega_i)}{m(\omega_i)}\right] \right] . \tag{11}$$

S is defined relative to a starting model, $m(\omega)$, for $A(\omega)$. It is termed the *default model*, because it is the image to

which the maximum-entropy method will default in the absence of data. It is usually chosen to be the smoothest function consistent with prior knowledge such as sum rules. If, before the quantum Monte Carlo simulation, we are completely ignorant about $A(\omega)$ except for the sum rule, we take $m(\omega)$ to be flat with a magnitude to satisfy the sum rule in the frequency range of interest. More generally, $m(\omega)$ should be chosen by the maximum-entropy method using the prior knowledge (e.g., moments) as constraints (or "testable information"[3]).

While space does not permit a detailed presentation of the variety of rationales leading to Eqs. (10) and (11), condensed-matter scientists will understand a statistical mechanics analogy. The thermodynamic entropy is the logarithm of the number of different states by which one can arrive at a given total energy, or any other macroscopic constraint. Analogously, the information theory entropy is the logarithm of the number of ways by which one can arrive at a given $A(\omega)$ in a Poisson process from the default model. For a flat $m(\omega)$, this is called the "monkey argument" by the image processing community[14]: "If a team of monkeys throws a very large number N of quanta randomly at the M a priori equivalent cells of an image, then the probability of obtaining a particular set $(n_1, n_2, \ldots, n_M)$ of occupation numbers shall be proportional to the degeneracy $N!/n_1!n_2! \cdots n_M!$." More generally, in a Poisson process the expected number of counts in a cell (or pixel) is $\bar{n}_i = am(\omega_i)\Delta\omega$. The parameter a characterizes the degree of fluctuation about the default model. Application of Stirling's formula to the factorials in a Poisson probability distribution then leads to Eqs. (10) and (11).

Finally, the image is obtained by maximizing the posterior probability. This is the same as functional variation of $A(\omega)$ to maximize the entropy, S, subject to a constraint on χ^2 with Lagrange multiplier $1/a$; hence the name maximum-entropy method. Clearly, $1/a$ controls the relative importance of the prior knowledge and the data. It is a statistical regularization parameter that is fixed by self-consistency arguments. Before the experiment the image starts at the default model and $a = \infty$. As the data improve a becomes smaller, and the image deviates from the default model (hopefully) toward the correct $A(\omega)$ acquiring sharper resolution. In traditional (or *historic*) ME, $1/a$ is chosen so that the image maximizes the entropy subject to it being "feasible," in that it fits the data according to $\chi^2 \leq N_\tau$, where N_τ is the number of independent data. This is, of course, consistent with the χ^2-probability distribution for independent Gaussian errors. In the more modern (or *classic*) ME,[15] $1/a$ is itself determined by Bayesian arguments that maximize the posterior probability of a given the data, $P[a|G_d(\tau)]$. The image exhibited is the center of a prob-

ability distribution (called the "bubble") of possible images. It is the classic ME formulation that makes possible estimates of the statistical significance of the features in the image.

Maximum entropy is a statistical inference principle, how one implements the principle is an algorithm, choosing the appropriate prior knowledge is physics, and incorporating it in the algorithm is an art. Prior knowledge should be built into the default model and into the data as a constraint. The images are conditional on the prior knowledge. The use of prior knowledge can be essential to solving some problems, but invalid prior knowledge can also lead to spurious results. For example, using a functional model for $A(\omega)$ defined by a few parameters is a strong form of incorporating prior knowledge. However, if the model cannot be fit to the data, at least some of the "prior knowledge" contained in the model is wrong. In many image-processing situations, one uses the flexibility of ME to iteratively input increasing amounts of prior knowledge (e.g., first positivity only, then include symmetry, and so on), or to test the validity of assumed "prior knowledge." Often, an initial ME image will suggest additional prior knowledge that should be incorporated. ME images often lead to a physical model for the quantity of interest, the parameters of which can then be estimated directly using the raw data.

IV. EXAMPLES OF ME ANALYTIC CONTINUATION

Let us consider the spectral function shown in Fig. 2(a) termed the "truth," where we have binned the spectral function into 41 pixels in the range $-\Omega \leq \omega - \mu \leq \Omega$. Using this in the transform, Eq. (2), with $\beta = 10.0$, $\Omega = 5.0$, binning with $\Delta\tau = 0.125$, and adding 1% relative Gaussian random noise yields the $G(\tau)$ data shown in Fig. 2(b) termed the "mock data." The problem then is to recover the spectral function from this mock data. Figure 2(c) shows the ME reconstruction of the spectral function using the classical ME algorithm. One can see that the basic structure of the truth is recovered in the reconstruction. However, the features at large $|\omega - \mu|$ are broadened compared to features at small $|\omega - \mu|$.

Most significantly, classic ME allows one to place error bars on the integrated intensities of the features in the image. The areas to be integrated over are marked by the hatching. The corresponding integrated intensity and estimated error are shown. The central peak and the gap for small $|\omega - \mu|$ are well determined, but the side peaks and the gap at large $|\omega - \mu|$ are less well determined.

We compare ME with the method of White *et al.* (WSSB).[9] Their method is to minimize a modified least-squares measure for the data which in our notation is

$$\chi^2_{\text{WSSB}} = \sum_k \frac{[G_d(\tau_k) - G_f(\tau_k)]^2}{\sigma_k^2} + b \sum_i [A(\omega_i) - A(\omega_{i+1})]^2 + h \sum_i \Theta(-A(\omega_i))A^2(\omega_i) \ .$$
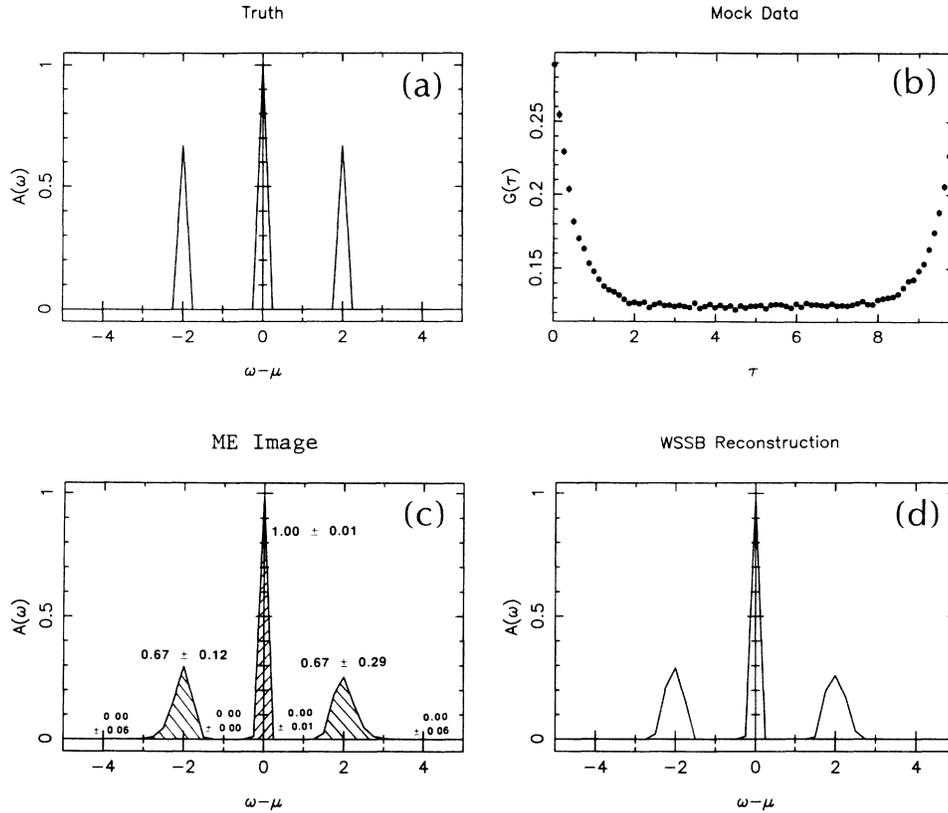
(12)

FIG. 2. First simulation, with parameters $\Omega=5$, $\Delta\omega=0.25$, $\beta=10$, and $\Delta\tau=0.125$: (a) true spectral function; (b) resultant mock data, $G(\tau)$, with relative Gaussian random error of 1%; (c) ME reconstruction; (d) reconstruction using the WSSB (Ref. 9) algorithm.

The first term is conventional. The second term with parameter b is designed to encourage smoothness and, hence, minimize the structure in the reconstruction. The third term with parameter h is designed to encourage (but not enforce) the positivity of the spectral function. The authors recommend how to iteratively adjust the parameters b and h to arrive at $\chi^2=1.1\chi^2_{min}$, where χ^2_{min} is the lowest value attained when enforcing positivity but very little smoothness. The resulting image is shown in Fig. 2(d). This image is essentially the same as the ME image, but it lacks information about the statistical reliability of the image. The WSSB approach is sensible since it incorporates the additional prior knowledge of positivity, but it is also *ad hoc* because there is no self-consistent way of choosing the parameters b and h. Moreover, the smoothness assumption (correlations between neighboring pixels) is invalid prior knowledge, as there are spectral functions that are not smooth (such as the Fano-Anderson model we consider later).

The pathology of the smoothness assumption can be shown with the spectral function in Fig. 3(a), which includes a sharp peak at $|\omega-\mu|=0$ and broad structure at large $|\omega-\mu|$. Figure 3(b) shows the corresponding mock data. Figure 3(c) shows the WSSB reconstruction at $\chi^2=1.1\chi^2_{min}$, which produces spurious structure. Figure 3(d) shows the image obtained by turning the smoothness up to achieve $\chi^2=2.0\chi^2_{min}$, and Fig. 3(e) has further smoothing with $\chi^2=10.0\chi^2_{min}$. One can see that the fits to the broad structure are improving at the expense of

broadening the central peak. Finally, Fig. 3(f) shows the ME reconstruction (symmetry was not enforced). It is much closer to the original image, and it permits estimates of the statistical reliability of features in the image, as shown.

The method of Jarrell and Biham[8] also attempts to enforce smoothness and positivity while working with data in Matsubara frequency space rather than imaginary time. There is no explicit provision for statistical error propagation in their method. It shows similar qualitative behavior to the WSSB method.

For a more realistic example of ME analytic continuation, we consider the spectral function of the Fano-Anderson model of an impurity state in a continuum. Here, the Hamiltonian is

$$\hat{H}=e_c\hat{b}^{\dagger}\hat{b}+\sum_k e_k\hat{a}^{\dagger}_k\hat{a}_k+\sum_k A_k(\hat{a}^{\dagger}_k\hat{b}+\hat{b}^{\dagger}\hat{a}_k) ,   (13)$$

where we consider a one-dimensional band with

$$A_k=\sqrt{c/L} ;   e_k=-w\cos k .   (14)$$

Here $\hat{b}^{\dagger}$ is the creation operator for an electron on the impurity site, $\hat{a}^{\dagger}_k$ is the creation operator for an electron in the band, e_c is the impurity level energy, $2w$ is the bandwidth, A_k is the coupling between the impurity level and the band, and L is the lattice spacing. For this model, one can analytically derive the spectral function for the impurity state from the imaginary-time impurity Green's function
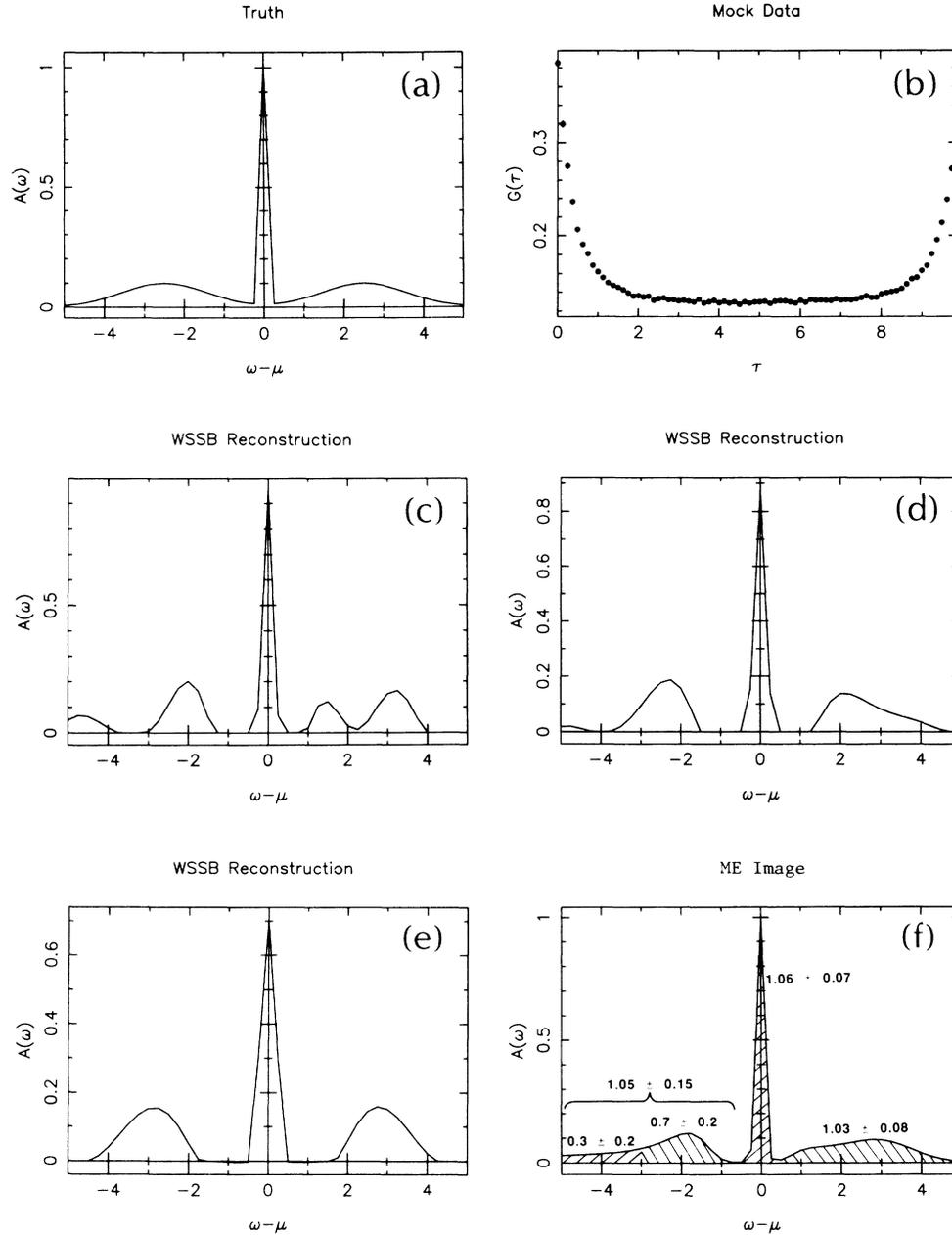
Truth

(a)

Mock Data

(b)

WSSB Reconstruction

(c)

WSSB Reconstruction

(d)

WSSB Reconstruction

(e)

ME Image

(f)

1.06 · 0.07

1.05 ± 0.15

0.7 ± 0.2

0.3 ± 0.2

1.03 ± 0.08

FIG. 3. Second simulation, with parameters $\Omega = 5$, $\Delta\omega = 0.25$, $\beta = 10$, and $\Delta\tau = 0.125$: (a) true spectral function; (b) resultant mock data $G(\tau)$ with relative Gaussian random error of 1%; (c) reconstruction using the WSSB algorithm; (d) WSSB reconstruction with extra smoothing, i.e., $\chi^2 = 2.0\chi^2_{min}$ instead of the recommended $\chi^2 = 1.1\chi^2_{min}$; (e) WSSB reconstruction with further smoothing, $\chi^2 = 10.0\chi^2_{min}$; (f) ME reconstruction.

$$G_b(\tau) \equiv -\langle T_\tau[\hat{b}(\tau)\hat{b}^\dagger(0)]\rangle .$$ (15)

For a variety of choice of parameters, we can generate mock $G_b(\tau)$ data by discretizing in τ and adding Gaussian random noise. We then examine the ability of ME to recover the impurity spectral function from such data.

Figure 4 shows the resulting simulations with parameters as indicated in the figures. In all cases the spectral function has a continuum between -1 and 1 and two δ function peaks on each side with total contributions Z_+

and Z_- to the sum rule, Eq. (4). The $G_b(\tau)$ data generated now has many more τ points and much greater statistical accuracy than in the simulations of Figs. 2 and 3 in order to provide sufficient information to reconstruct the more complex spectral function. These errors are smaller than those used in Refs. 8, 9, and 11, but they are within the range of the current state of the art in quantum Monte Carlo simulations. The data shown in the middle of the figure are clearly insensitive to the detailed structure in the spectral function, which is what makes this
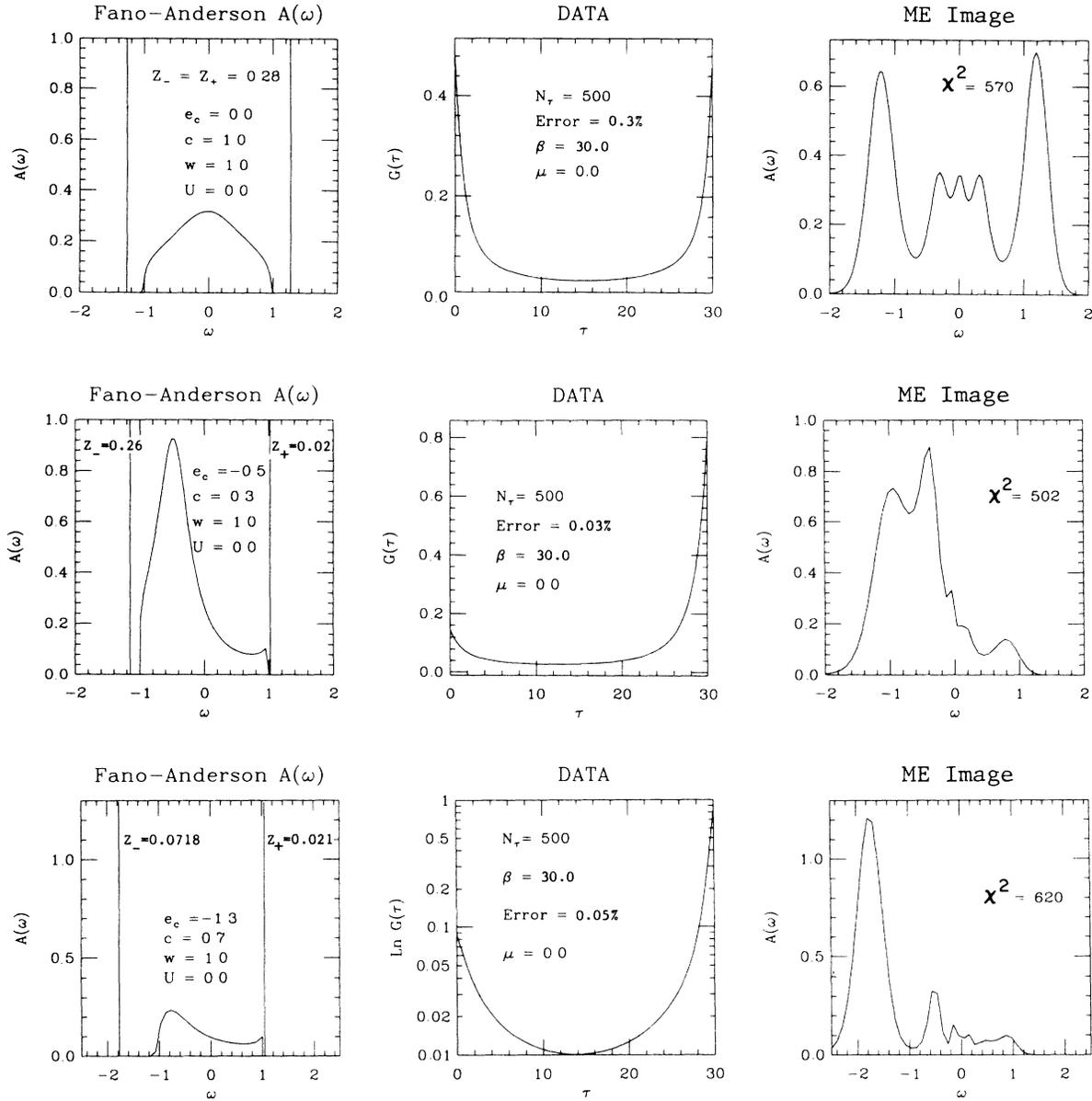
FIG. 4. ME simulations for the Fano-Anderson model of an impurity state in a continuum. In each case, the left-most figure is the spectral function for the indicated choice of parameters, the middle figure are mock data for $G(\tau)$, and the right-most figure is the ME reconstruction. The total area under the spectral function is one. The continuum is always chosen to extend from -1.0 to 1.0, but the position of the impurity level and its coupling to the continuum is varied. Discrete states are split off from the continuum that contribute δ functions to the spectral functions (indicated by the solid vertical lines) of weights Z_+ and Z_-. N_τ is the number of data points of $G(\tau)$; *Error* is the fractional standard deviation of the error in $G(\tau)$; β is the inverse temperature taken to be 30.0; and μ is the chemical potential taken to be 0.0. The χ^2 is the actual χ^2 achieved in the ME reconstruction, to be compared with the N_τ. Qualitative and semiquantitative features of the true spectral function are reproduced by the ME reconstruction, but there is a tendency toward broadening at large $|\omega|$ and toward cuspy structure at small $|\omega|$. The error analysis of the classic ME method indicates that the cuspy structure is not statistically significant.

problem extremely ill posed. Nevertheless, one can see that the ME images on the right of the figure have recovered the essential features of the original spectral functions shown on the left. However, there is a tendency to put cuspy structure in the image for small $|\omega-\mu|$, and to broaden structure at large $|\omega-\mu|$. This tendency is a feature of any image-reconstruction method, as will be explained in the following section. The cuspy struc-

ture at small $|\omega-\mu|$ can be shown to be spurious by again using the classical ME option of estimating the statistical errors on the integrated intensity of that structure.

We conclude that ME appears to be a very powerful image-reconstruction method with which to address the analytic continuation problem for quantum Monte Carlo data.

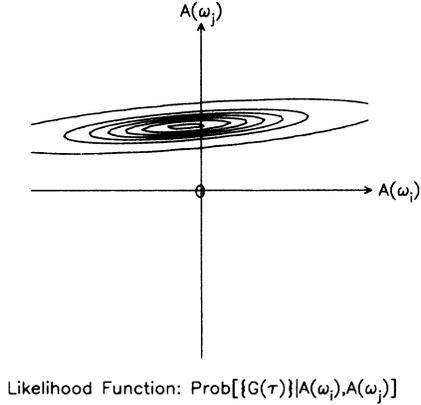Likelihood Function: Prob[{G(τ)}|A(ωᵢ),A(ωⱼ)]

FIG. 5. A section through a schematic likelihood function. The data constrain the distribution well in some directions but poorly in others. The good directions are associated with large eigenvalues of the Hessian matrix (log-likelihood function) and the bad directions with small ones.

V. THE LIKELIHOOD FUNCTION OF THE ANALYTIC CONTINUATION PROBLEM

In this section, we address error propagation features of the analytic continuation problem that are independent of the choice of image-reconstruction method. These features are embodied in the likelihood function given by Eqs. (7)–(9).

The image that maximizes the likelihood function, $A_M(\omega)$, can be obtained by setting to zero the functional derivative of the likelihood function with respect to the pixels, $(\mathbf{A})_i \equiv A(\omega_i)$. That is, $A_M(\omega)$ is the solution of $\nabla_A P[G_d|A] = 0$. Then the likelihood function contains information about the probability of fluctuations about this maximum

$$(\delta \mathbf{A})_i \equiv A(\omega_i) - A_M(\omega_i) \ .$$

It can be written as

$$P[G_d(\tau)|A(\omega)] \propto \exp(-\tfrac{1}{4}\delta \mathbf{A}^T \cdot \overleftrightarrow{\mathbf{H}} \cdot \delta \mathbf{A}) \ , \tag{16}$$

where the Hessian matrix is

$$(\overleftrightarrow{\mathbf{H}})_{ij} = \frac{\partial^2 \chi^2}{\partial A(\omega_i) \partial A(\omega_j)} \ . \tag{17}$$

Since the eigenvectors of $\overleftrightarrow{\mathbf{H}}$ form a complete set, the true spectral function can be decomposed into a linear sum of eigenvectors. Only the coefficients of eigenvectors with large eigenvalues can be reliably determined from the data. Eigenvectors associated with small eigenvalues correspond to properties of the spectral function about which there is a lack of evidence in the data. A schematic likelihood function is illustrated in Fig. 5, showing good directions and bad directions. Any successful image-reconstruction technique must recover the components of the spectral function along the good directions while suppressing the components along the bad directions. (Some readers may note that this is the basis of image reconstruction by the method of singular value decomposition.)

The open circles in Fig. 6 shows specific calculations of the eigenvalue spectrum of the Hessian matrix (or log-likelihood function) for the analytic continuation problem. The parameters are as in Fig. 2, but the statistical errors are approximated as constant to emphasize features of the analytic continuation problem that are independent of the particular spectral function under consideration. One can see that the eigenvalue spectrum is a very steep function, so that the only credible information in the data is the few components of the spectral function along the good eigenvectors. Figures 7(a)–7(d) show some of the eigenvectors. Good eigenvectors have cuspy structure for small $|\omega - \mu|$ and only broad structure for large $|\omega - \mu|$. Sharp structure at large $|\omega - \mu|$ corresponds to bad eigenvectors about which there is little information in the data. These are suppressed by ME or any other valid image-reconstruction method. Just such tendencies are seen in the images of the Fano-Anderson model shown in Fig. 4.

One can use these calculations to estimate the sensitivity of the reconstruction to the parameters of the quantum Monte Carlo simulations, which can aid in the design of simulations. The triangles in Fig. 6 shows changes to the eigenvalue spectrum when the Monte Carlo parameters are changed. The open circles correspond to the same parameters as in Fig. 2. The altered parameters are as follows: Fig. 6(a) decreasing the statistical noise from 0.01 to 0.001; Fig. 6(b) decreasing the time interval from $\Delta\tau = 0.125$; Fig. 6(c) increasing the temperature from $\beta = 10$ to 2.5. In Figs. 6(a) and 6(b), the eigenvalue spectrum retains its shape and simply scales with the changes in parameters corresponding to the increasing amount of data available. However, in Fig. 6(c) the eigenvalue spectrum changes its shape with changing temperature, indicating that lower temperatures have more information about the spectral function.

These calculations have an extremely important caveat: Are the data obtained by quantum Monte Carlo simulations in fact statistically independent? As far as we know, this issue has not been addressed either theoretically or numerically. We suspect that there is nothing in the logic of the quantum Monte Carlo method to assure statistical independence. This may not be a serious consideration if all one wants is a "pretty picture" of a possible spectral function that fits the data. A more careful Monte Carlo calculation would provide not merely the errors, σ_k, but the full covariance matrix, C_{ij}. Clearly, one has $C_{kk} = \sigma_k^2$. The statistical independence assumption fails if C_{ij} has significant off-diagonal components. The obvious generalization of the likelihood function would be

$$P[G_d|A] \propto \exp(-\tfrac{1}{2}\delta \mathbf{G}^T \cdot \overleftrightarrow{\mathbf{C}}^{-1} \cdot \delta \mathbf{G}) \ ,$$

$$(\delta \mathbf{G})_i \equiv G_d(\tau_i) - G_f(\tau_i) \ , \tag{18}$$

where $\overleftrightarrow{\mathbf{C}}^{-1}$ is the inverse of the covariance matrix. Significant covariance in the quantum Monte Carlo data will complicate the algebra of any image reconstruction method, but it is essential to propagate statistical errors correctly. In the classical ME method, for example, ignoring significant covariance would limit the ability to

Eigenvalue Spectrum (decreasing noise)

Eigenvalue Spectrum (decreasing Δτ)
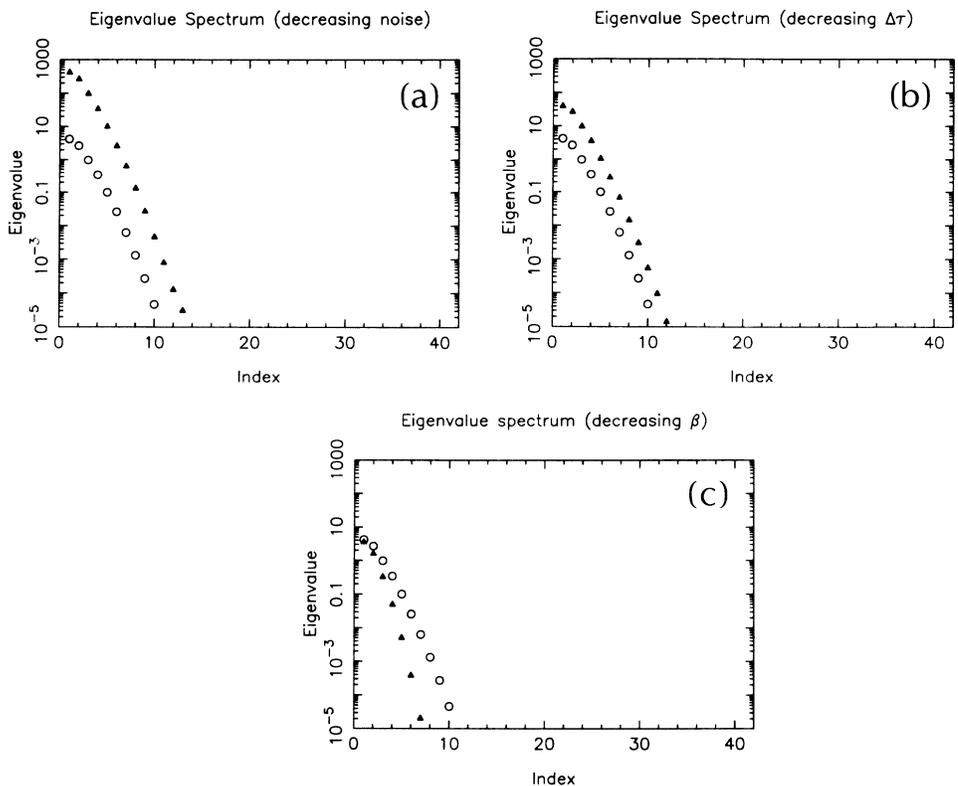
Eigenvalue spectrum (decreasing β)

FIG. 6. Eigenvalues of the Hessian matrix (log-likelihood function) for the analytic continuation problem, indexed in order of decreasing eigenvalue. The open circles in (a)–(c) have the same parameters: $\Delta\omega=0.25$, $\beta=10$, $\Delta\tau=0.125$, and $\sigma=0.01$. The triangles correspond to changes in these parameters: (a) decreasing statistical noise to $\sigma=0.001$; (b) decreasing the time interval, $\Delta\tau=0.0125$; (c) increasing the temperature, $\beta=2.5$.

Eigenvectors: 1 – 3

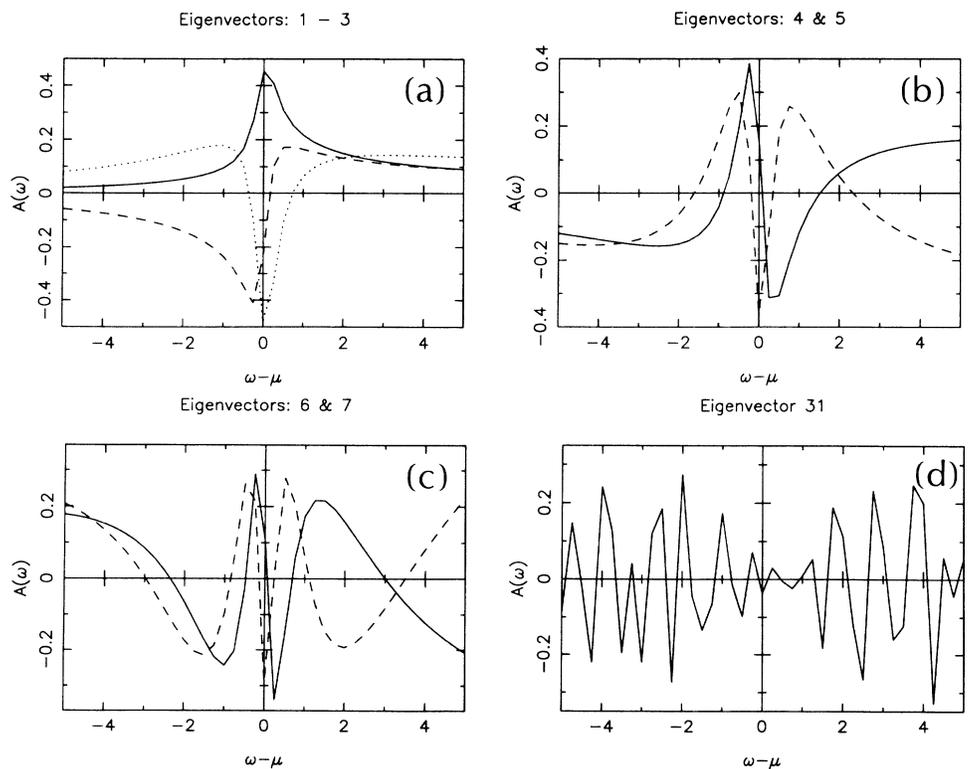Eigenvectors: 4 & 5

Eigenvectors: 6 & 7

Eigenvector 31

FIG. 7. Eigenvectors of the Hessian matrix (log-likelihood function) for the analytic continuation problem. Parameters are: $\Delta\omega=0.25$, $\beta=10$, and $\Delta\tau=0.125$ and an absolute statistical accuracy for $G(\tau)$ of $\sigma=$const$=1.0$. The relations between eigenvectors and eigenvalue index are: (a) solid 1, dashed 2, dotted 3; (b) solid 4, dashed 5; (c) solid 6, dashed 7; (d) solid 31.

choose the appropriate weighting between the data and the prior knowledge (or $1/a$), or to estimate the statistical reliability of the integrated intensities of features in the spectral function.

The covariance matrix is real and symmetric. Therefore, in principle, it can be diagonalized by an orthogonal (unitary) transformation in the data space, and it should have only positive eigenvalues. This should allow the application of existing image-reconstruction algorithms to the quantum Monte Carlo data, albeit in a transformed data space. The nonindependence of the data will result in many large eigenvalues of the covariance matrix that do not significantly constrain the spectral function, and the corresponding eigenvectors may be ignored. This procedure might be implemented, for example, by singular value decomposition of the covariance matrix.

VI. CONCLUSIONS

We have shown[16] that the analytic continuation of imaginary-time quantum Monte Carlo data to obtain real-frequency spectral functions can be addressed as an image-reconstruction problem. We have described and illustrated the use of the maximum-entropy method for this problem, and we have reported encouraging prelimi-

nary results for the spectral function of the Fano-Anderson model of an impurity state coupled to a continuum. We have compared our results to other proposals for analytic continuation in the literature. We have also addressed issues of statistical error propagation in the analytic continuation of quantum Monte Carlo data, which are independent of the choice of image-reconstruction technique.

The success of the maximum-entropy method for the analytic continuation problem opens the door to studying the dynamical properties of quantum many-body systems by quantum Monte Carlo methods.[17] What remains is to resolve the issues of statistical error propagation in Monte Carlo calculations, and to begin to apply ME to real data.

*Present address: MS B262, Theoretical Division and Los Alamos Neutron Scattering Center.

†Present address: MS B262, Theoretical Division.

[1]For a tutorial review of fermion Monte Carlo methods, see E. Loh and J. E. Gubernatis, in *Electronic Phase Transitions,* edited by W. Hanke and Y. Kopaev (Elsevier, New York, 1990).

[2]F. S. Acton, *Numerical Methods that Work* (Harper and Row, New York, 1970), pp. 245–257.

[3]E. T. Jaynes, in *The Maximum Entropy Formalism,* edited by R. D. Levine and M. Tribus (MIT Press, Cambridge, 1978), pp. 15–118.

[4]MEMSYS 3, produced by Maximum Entropy Data Consultants Ltd. (1989); this is an implementation of the Cambridge algorithm. For details of MEMSYS 1, see J. Skilling and R. K. Bryan, Mon. Not. R. Astron. Soc. **211**, 114 (1984).

[5]H. Jeffreys, *Theory of Probability* (Oxford University Press, London, 1939); R. P. Cox, Am. J. Phys. **17**, 1 (1946).

[6]S. F. Gull and J. Skilling, IEEE Proc. **131(F)**, 646 (1984); R. Narayan and R. Nityananda, Annu. Rev. Astrophys. **24**, 127 (1986).

[7]G. D. Mahan, *Many Particle Physics* (Plenum, New York,

1981), p. 256ff.

[8]M. Jarrell and O. Biham, Phys. Rev. Lett. **63**, 2504 (1989).

[9]S. R. White, D. J. Scalapino, R. L. Sugar, and N. E. Bickers, Phys. Rev. Lett. **63**, 1523 (1989).

[10]H. B. Schuttler and D. J. Scalapino, Phys. Rev. Lett. **55**, 1204 (1985); Phys. Rev. B **34**, 4744 (1986).

[11]J. Hirsch, in *Quantum Monte Carlo Methods,* edited by M. Suzuki (Springer-Verlag, Heidelberg, 1987).

[12]J. E. Shore and R. W. Johnson, IEEE Trans. Inf. Theory **IT-26**, 26 (1980); **IT-29**, 942 (1980); Y. Tikochinsky, N. Z. Tishby, and R. D. Levine, Phys. Rev. Lett. **52**, 1357 (1984).
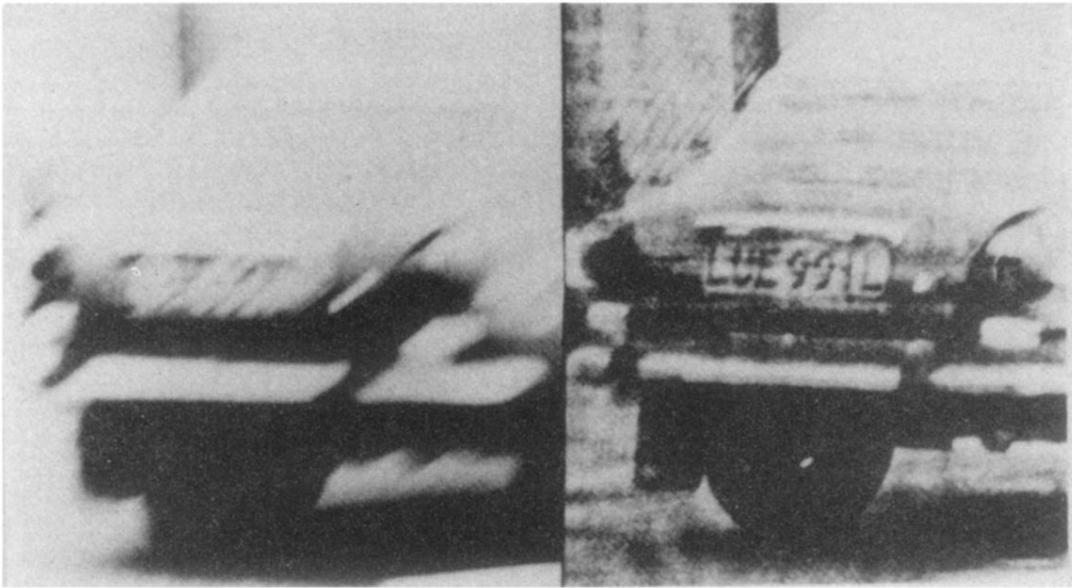
[13]C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).

[14]S. F. Gull and G. J. Daniell, Nature **272**, 686 (1978).

[15]J. Skilling, in *Maximum Entropy and Bayesian Methods,* edited by J. Skilling (Kluwer Academic Publishers, 1989).

[16]A preliminary account of this work by R. N. Silver, D. S. Sivia, and S. E. Gubernatis appeared in *Proceedings of the Quantum Simulations of Condensed Matter Phenomena,* edited by J. Gubernatis (World Scientific, Singapore, in press).

[17]See also the contributions of K. Haug *et al.* in Ref. 16.

Data   Reconstruction

FIG. 1. Optical deconvolution: an example of the use of ME.