Analytic continuation of quantum Monte Carlo data by stochastic analytical inference

Sebastian Fuchs* and Thomas Pruschke
Institut für Theoretische Physik, Georg-August-Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

Mark Jarrell
Louisiana State University, Baton Rouge, Louisiana 70803, USA

(Received 3 January 2010; revised manuscript received 11 March 2010; published 4 May 2010)

We present an algorithm for the analytic continuation of imaginary-time quantum Monte Carlo data which is strictly based on principles of Bayesian statistical inference. Within this framework we are able to obtain an explicit expression for the calculation of a weighted average over possible energy spectra, which can be evaluated by standard Monte Carlo simulations, yielding as by-product also the distribution function as function of the regularization parameter. Our algorithm thus avoids the usual *ad hoc* assumptions introduced in similar algorithms to fix the regularization parameter. We apply the algorithm to imaginary-time quantum Monte Carlo data and compare the resulting energy spectra with those from a standard maximum-entropy calculation.

DOI: 10.1103/PhysRevE.81.056701 PACS number(s): 02.70.−c, 71.15.Dx, 71.28.+d, 71.10.Fd

I. INTRODUCTION

Quantum Monte Carlo (QMC) simulations are a powerful computational tool to calculate properties of interacting quantum many-particle systems, such as spin models or strongly correlated electron systems. Of particular interest in those systems are dynamical correlation functions such as single-particle spectra or susceptibilities, respectively, dynamical structure factors. However, QMC presently provides data only on the imaginary-time axis, and the necessary analytic continuation of these data has proven to be difficult.

The standard tool to solve this problem is the maximum-entropy method (MEM) [1]. It uses arguments of Bayesian logic [2,3] to obtain the most probable energy spectrum. In order to solve this optimization problem efficiently, the maximum entropy method approximates all occurring probability distributions to be of a Gaussian shape.

In the past efforts were made to provide an alternative to this approach [4–6]. It was proposed to perform a Monte Carlo average over a wide range of spectra instead of selecting a single spectrum. So far, the method lacked a rigorous rule to eliminate a regularization parameter inherent in the algorithm. Although this approach has been interpreted in terms of Bayesian inference [6], none of the authors utilized Bayesian logic to eliminate the regularization parameter.

We show that this stochastic approach can also be understood in terms of Bayesian statistical inference. We derive a strict criterion to eliminate the free parameter, which is completely based on Bayesian logic. It uses Monte Carlo techniques to both calculate the average spectrum and to eliminate the regularization parameter. It treats all probabilities exactly and hereby avoids the approximations made in the maximum entropy method. We apply the algorithm to imaginary-frequency quantum Monte Carlo data and compare the resulting spectra with results from maximum entropy calculations.

*fuchs@theorie.physik.uni-goettingen.de

II. PROBLEM OF ANALYTIC CONTINUATION

For a finite temperature T quantum Monte Carlo simulations can provide accurate estimates $\bar{G}_n$ for either imaginary-time correlation function $G(\tau)$ at a finite set of N imaginary-time points τ_n or, alternatively, for imaginary-frequency correlation functions $G(i\omega_n)$ at a finite set of N Matsubara frequencies ω_n. The frequencies are defined as $\omega_n=(2n+1)\pi/\beta$ for fermions and as $\omega_n=2n\pi/\beta$ for bosons with $\beta=1/k_{\mathrm{B}}T$.

Because of the stochastic nature of Monte Carlo algorithms each of the $\bar{G}_n$ possesses a known statistical error. Moreover, the data for the different time or frequency points are usually highly correlated. Therefore the input to the analytic continuation procedure consists of the Monte Carlo estimates $\bar{G}_i$ and their covariance matrix

$$C_{nm}=\overline{G_nG_m}-\bar{G}_n\bar{G}_m. \tag{1}$$

In principle the spectral function $A(\omega)=-\frac{1}{\pi}\mathrm{Im}\,G(\omega+i0^+)$ can be extracted from these data by inverting

$$G(\tau_n)=\int d\omega\, K_n(\omega)A(\omega) \tag{2}$$

with

$$K_n(\omega)=K(\tau_n,\omega):=-\frac{e^{-\omega\tau}}{1\pm e^{-\omega\beta}} \tag{3}$$

for time-dependent data or

$$K_n(\omega)=K(i\omega_n,\omega):=\pm\frac{1}{i\omega_n-\omega} \tag{4}$$

for frequency-dependent data, where the upper sign holds for fermions and the lower one for bosons. The spectral function is normalized to

$$\mathcal{N} = \int d\omega\, A(\omega) \tag{5}$$

and is non-negative for all ω. However, a direct inversion of Eq. (2) is an ill-posed problem and numerically impossible.

A least-squares fit of $A(\omega)$ to the data $\bar{G}_n$ minimizes the χ^2 *estimate*

$$\chi^2[A] = \sum_{n,m} (\bar{G}_n - G(\tau_n))^* \sqrt{C^{-1}_{nm}} (\bar{G}_m - G(\tau_m)) \tag{6}$$

with respect to $A(\omega)$. This approach leads to a multitude of different solutions and consequently cannot solve the problem either.

A. Maximum-entropy method

The maximum-entropy method can be understood as an attempt to regularize the least-squares fit described above. One defines the entropy

$$S[A] = -\int d\omega\, A(\omega) \ln \frac{A(\omega)}{D(\omega)} \tag{7}$$

relative to a *default model* $D(\omega)$. Any information, which is known about the spectrum beforehand, can be encoded in the default model. If $D(\omega)$ is non-negative and possesses the same norm $\mathcal{N}$ as the spectrum $A(\omega)$, the entropy S will be nonpositive and maximal for $D(\omega)$. Instead of just minimizing χ^2 the MEM minimizes the quantity

$$Q[A] = \frac{1}{2}\chi^2[A] - \alpha S[A] \tag{8}$$

introducing a *regularization parameter* α. This optimization problem can be numerically solved for fixed α to find the minimizing spectrum $\hat{A}_\alpha(\omega)$. In the limit of $\alpha \to \infty$ the spectrum minimizing Q is the default model $D(\omega)$. For $\alpha \to 0$ the least-squares fit is regained. Thus the parameter α interpolates between the fit result and the default model.

In order to find a criterion to eliminate the parameter different approaches exist. The simplest rule is to take the spectrum where $\chi^2 \sim N$. This choice ensures that the differences between model and data are of the order of the error bars thereby avoiding overfitting. In order to derive more sophisticated methods the MEM needs to be reinterpreted by means of Bayesian statistical inference [2,3].

1. Bayesian statistical inference

The MEM can be reformulated by defining subjective probabilities for the quantities involved in the analytic continuation problem. Let $P[A]$ denote the *prior probability* of the spectrum $A(\omega)$. $P[A|\bar{G}]$ denotes the *posterior probability* of A given the input data $\bar{G}$ and $P[\bar{G}|A]$ the *likelihood function*. *Bayes's Theorem* [7] relates these probabilities to each other:

$$P[A|\bar{G}] = P[\bar{G}|A]P[A]/P[\bar{G}]. \tag{9}$$

The probability $P[\bar{G}]$ is called the *evidence* and serves as normalization for the posterior probability $P[A|\bar{G}]$,

$$P[\bar{G}] = \int \mathcal{D}A\, P[\bar{G}|A]P[A]. \tag{10}$$

One identifies

$$P[\bar{G}|A] = \frac{1}{Z_1}\exp\left(-\frac{1}{2}\chi^2[A]\right) \tag{11}$$

and

$$P[A] = \frac{1}{Z_2}\exp(\alpha S[A]). \tag{12}$$

The quantities

$$Z_1 = \int \mathcal{D}\bar{G}\, e^{-\chi^2[A]/2} \tag{13}$$

and

$$Z_2 = \int \mathcal{D}A\, e^{\alpha S[A]} \tag{14}$$

normalize the respective probabilities. This way the posterior probability can be rewritten as

$$P[A|\bar{G}] = \frac{e^{-Q[A]}}{Z_1 Z_2 P[\bar{G}]}. \tag{15}$$

with

$$P[\bar{G}] = \frac{\int \mathcal{D}A\, e^{-Q[A]}}{Z_1 Z_2} \tag{16}$$

Thus the minimization of Q can be reinterpreted as the maximization of the posterior probability $P[A|\bar{G}] \sim e^{-Q}$. The MEM therefore determines the *most probable* spectrum $\hat{A}_\alpha$ given the input data $\bar{G}$.

2. Bayesian inference and the regularization parameter α

This alternative formulation of the problem provides the necessary tools to eliminate the free parameter α [8,9]. Equation (9) can be rewritten including α,

$$P[A,\alpha|\bar{G}] = P[\bar{G}|A,\alpha]P[A,\alpha]/P[\bar{G}]. \tag{17}$$

If one applies Bayes's theorem to factorize $P[A,\alpha]$ and integrates over A, the relation

$$\begin{aligned} P[\alpha|\bar{G}] &= P[\alpha]\int \mathcal{D}A\, P[\bar{G}|A,\alpha]P[A|\alpha]/P[\bar{G}] \\ &= \frac{P[\alpha]}{Z_1 Z_2 P[\bar{G}]}\int \mathcal{D}A\, e^{-Q[A]} \end{aligned} \tag{18}$$

for the posterior probability $P[\alpha|\bar{G}]$ can be found. Analogous to the argument given above, one identifies $P[\bar{G}|A,\alpha] \sim \exp(-\frac{1}{2}\chi^2[A])$ and $P[A|\alpha] \sim \exp(\alpha S[A])$. The evidence

$$P[\bar{G}] = \int d\alpha \frac{P[\alpha] \int \mathcal{D}A \, e^{-Q[A]}}{Z_1 Z_2} \tag{19}$$

is an α-independent normalization constant. All quantities in this equation are known except $P[\alpha]$, the prior probability of α. It is either taken to be constant or to be the Jeffreys prior $1/\alpha$ [9–11]. However, the choice of $P[\alpha]$ turns out to be of little influence on the resulting spectra.

By assuming all probabilities involved to be of a Gaussian shape a numerical treatment of Eqs. (15) and (18) is possible. There are two alternatives:

(1) one calculates α^* as the α that maximizes $P[\alpha|\bar{G}]$ and takes $\hat{A}_{\alpha^*}$ as the final result for the spectral function [8,9];

(2) one averages over all $\hat{A}_\alpha$ weighted by the posterior probability of α, i.e., the average spectrum

$$\langle A \rangle = \int d\alpha \, P[\alpha|\bar{G}] \hat{A}_\alpha, \tag{20}$$

(3) is taken as the final result [11].

It is not *a priori* clear, which of the two algorithms is favorable.

B. Stochastic analytical inference

Stochastic analytical inference (SAI) is an alternative to the standard MEM which does not employ the explicit regularization of the fit by entropy Eq. (7). Rather than maximizing $P[A|\bar{G}]$ an average over all possible spectra weighted by

$$w \sim \exp\left(-\frac{1}{2}\chi^2/\alpha\right) \tag{21}$$

is performed. Beach refined this approach, by introducing the default model $D(\omega)$ of the MEM into the algorithm [5]. By mapping ω unto $x \in [0,1]$ using

$$x = \phi(\omega) = \frac{1}{N}\int_{-\infty}^{\omega} d\omega' D(\omega'), \tag{22}$$

a dimensionless field $n(x)$ can be defined as

$$n(x) = \frac{A(\phi^{-1}(x))}{D(\phi^{-1}(x))}. \tag{23}$$

The field $n(x)$ is normalized to 1,

$$1 = \int_0^1 dx n(x). \tag{24}$$

By calculating the average field,

$$\langle n(x) \rangle_\alpha = \frac{1}{Z}\int \mathcal{D}'n(x) \, n(x) \, e^{-\chi^2[n(x)]/2\alpha} \tag{25}$$

with

$$Z = \int \mathcal{D}'n(x) \, e^{-\chi^2[n(x)]/2\alpha}. \tag{26}$$

The measure

$$\mathcal{D}'n(x) = \mathcal{D}n(x) \, \Theta[n] \, \delta\left(\int_0^1 dx n(x) - 1\right) \tag{27}$$

restricts the integration to fields $n(x)$ that satisfy norm rule Eq. (24) and the positivity requirement. In Eq. (27),

$$\Theta[n] = \begin{cases} 1, & \text{if } \forall x\text{: } n(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

The average spectrum $\langle A \rangle_\alpha$ can be regained via

$$\langle A(\omega) \rangle_\alpha = D(\omega) \langle n(\phi(\omega)) \rangle_\beta. \tag{28}$$

If $\frac{1}{2}\chi^2$ is interpreted as a Hamiltonian of a fictitious physical system, Eq. (25) possesses the structure of a canonical ensemble average at a temperature α. The laws of statistical mechanics then state that the average spectral function $\langle A \rangle_\alpha$ minimizes the free energy

$$F = \frac{1}{2}\langle \chi^2 \rangle_\alpha - \alpha \mathcal{S}. \tag{29}$$

This expression displays a similar structure as Eq. (8). Thus the averaging process implicitly generates an entropy $\mathcal{S}$. However, this entropy does not have the explicit form of Eq. (7). In the limit $\alpha \to 0$ the averaging process minimizes χ^2. Whereas in the limit $\alpha \to \infty$ the average in Eq. (25) is completely unaffected by χ^2 and will—constrained by Eq. (24)—result in $\langle n(x) \rangle = 1$. In this case the resulting spectrum is the default model. The algorithm therefore exhibits the same limiting cases as the MEM. Additionally, Beach showed that a mean-field treatment of the fictitious physical system described by χ^2 is formally equivalent to the MEM [5].

The remaining open question, namely, how to eliminate the parameter α, was addressed by all preceding authors differently:

(1) Sandvik proposed to examine the plot of the average entropy against α and identifies the final α by a sharp drop in the entropy curve [4];

(2) Beach examined a double-logarithmic plot of the average χ^2 and identifies the final α by a kink in the χ^2 curve [5]; and

(3) Syljuåsen argues to take $\alpha = 1$ [6].

All criteria are merely based on heuristic arguments. The simple rule to take $\chi^2 \sim N$ is also applicable to this method and should be mentioned here.

1. Bayesian statistical inference

In the following we will use Bayesian inference to derive a criterion to eliminate the regularization parameter α. In contrast to the MEM the stochastic analytical inference does not maximize the posterior probability $P[A|\bar{G}]$. Instead, it averages all possible fields n (omitting the argument x in the progress) weighted by $P[n|\bar{G}]$,

$$\langle n \rangle = \int \mathcal{D}n \, n \, P[n|\bar{G}]. \tag{30}$$

Bayes's theorem can be applied to factorize $P[n|\bar{G}]$ analogous to Eq. (9),

$$P[n|\bar{G}] = P[\bar{G}|n]P[n]/P[\bar{G}]. \tag{31}$$

SAI does not introduce an explicit entropy term. Following Ref. [6] only the positivity requirement and norm rule Eq. (24) enter the prior probability

$$P[n] = \Theta[n]\delta\left(\int_0^1 dx\, n(x) - 1\right). \tag{32}$$

The likelihood function is identified as

$$P[\bar{G}|n] = \frac{1}{Z'} e^{-\chi^2/2\alpha}. \tag{33}$$

By evaluating a Gaussian integral the normalization Z' can be readily calculated to be

$$Z' = \int \mathcal{D}\bar{G}\, e^{-\chi^2/2\alpha} = (2\pi\alpha)^{N/2}\sqrt{\det C}. \tag{34}$$

Using

$$P[\bar{G}] = \int \mathcal{D}' n \frac{e^{-\chi^2[n]/2\alpha}}{Z'} = \frac{Z}{Z'} \tag{35}$$

the posterior probability results in

$$P[n|\bar{G}] = \Theta[n]\delta\left(\int_0^1 dx\, n(x) - 1\right)\frac{1}{Z} e^{-\chi^2[n]/2\alpha}, \tag{36}$$

as expected from the comparison of the Eqs. (25) and (30).

2. Bayesian inference and the regularization parameter α

Bayesian logic can also be utilized to calculate the posterior probability $P[\alpha|\bar{G}]$. Substituting n for A in Eq. (18) and identifying $P[n|\alpha]=P[n]$ with $P[n]$ from Eq. (32) and correspondingly $P[\bar{G}|n,\alpha]=P[\bar{G}|n]$ with $P[\bar{G}|n]$ from Eq. (33), one obtains

$$P[\alpha|\bar{G}] = P[\alpha]\int \mathcal{D}n\, P[\bar{G}|n,\alpha]P[n|\alpha]/P[\bar{G}]$$

$$= \frac{P[\alpha]}{Z' P[\bar{G}]}\int \mathcal{D}' n\, e^{-\chi^2[n]/2\alpha}. \tag{37}$$

The evidence

$$P[\bar{G}] = \int d\alpha \frac{P[\alpha]e^{-\chi^2/2\alpha}}{Z'(\alpha)} \tag{38}$$

is again an α-independent normalization constant. The combination of Eqs. (34) and (37) gives the final expression for the α-dependence of the posterior probability,

$$P[\alpha|\bar{G}] \sim P[\alpha]\alpha^{-N/2}\int \mathcal{D}n\, e^{-\chi^2[n]/2\alpha}. \tag{39}$$

Analogous to the MEM one has two possibilities to treat the regularization parameter:

(1) one calculates α^* as the α that maximizes $P[\alpha|\bar{G}]$ and takes $\langle n\rangle_{\alpha^*}$ as the final result;

(2) one averages over all $\langle n\rangle_\alpha$ weighted by the posterior probability of α, i.e., the average field

$$\langle\langle n\rangle\rangle = \int d\alpha\, P[\alpha|\bar{G}]\langle n\rangle_\alpha \tag{40}$$

is taken as the final result.

III. MONTE CARLO EVALUATION

A. Configuration and update scheme

In order to calculate the quantities appearing in Eqs. (25) and (39) a numerically treatable approximation for the field configuration $n(x)$ and the integration measure $\mathcal{D}n$ has to be found. Our implementation closely follows Ref. [5]. The field configuration is represented by a superposition of delta function walkers with residues r_n and coordinates x_n,

$$n(x) = \sum_n r_n \delta(x - x_n). \tag{41}$$

The Monte Carlo updates consist of randomly proposed shifts of the coordinates x_n and random redistributions of the residues r_n. Redistributions that are not only norm conserving but also conserve higher moments of the configuration [4,5] were also implemented. Both the higher moment conserving updates and the coordinate shifts of the walkers have proven to be very effective in speeding up the algorithm.

Average Eq. (25) is evaluated by a standard Monte Carlo simulation using Metropolis weights. The regularization parameter α is treated as the temperature of the system. The simulation is performed for a wide range of different α-values. A parallel tempering [12–14] algorithm is necessary to ensure convergence for small α. In order to measure the average field configuration a histogram of the delta function walkers is recorded.

B. Calculation of the probability $P[\alpha|\bar{G}]$

A particular problem in the proposed approach is that a numerical treatment of Eq. (39) involves the calculation of the quantity

$$Z = \int \mathcal{D}' n\, e^{-\chi^2/2\alpha}. \tag{42}$$

This is equivalent to calculating a partition function in a canonical ensemble at temperature α. Standard Monte Carlo techniques are only able to calculate thermal expectation values but not the partition function itself. We use a *Wang-Landau algorithm* [15,16] to generate the density of states $\rho(E)$ of the system. Once $\rho(E)$ is calculated, the partition function can be obtained by

$$Z = \int dE\, \rho(E) e^{-E/\alpha}. \tag{43}$$

The Wang-Landau algorithm performs a random walk in energy space with probability $p(E)=1/\rho(E)$ using the usual metropolis weights. Since the density of states is unknown at the beginning of the simulation, one starts with an arbitrary

starting value, e.g., $\rho(E)=1$. For each visited energy one updates an energy histogram and multiplies the density of states by an modification factor $f>1$. When the histogram is reasonably flat, one resets the histogram and restarts the simulation with a new modification factor $f'=\sqrt{f}$. The starting value of f is usually taken to be Euler's constant and the procedure is repeated until f is very close to 1 (16 times in our implementation). The resulting $\rho(E)$ is the density of states of the system up to an unknown normalization factor. In order to speed up the convergence of the algorithm, it is advisable to divide the energy range of interest into several slightly overlapping smaller intervals.

IV. SIMULATION RESULTS

A. Model

We apply the algorithm to imaginary-time data from quantum Monte Carlo simulations. As test case we consider the two-dimensional single-band Hubbard model

$$H=-t\sum_{\langle i,j\rangle\sigma}c_{i\sigma}^{\dagger}c_{j\sigma}+U\sum_{i}n_{i\uparrow}n_{i\downarrow}. \tag{44}$$

Here i and j are lattice site indices, the operators $c_{i\sigma}^{\dagger}$ ($c_{i\sigma}$) create (destroy) an electron with spin $\sigma\in\{\uparrow,\downarrow\}$ at site i, $n_{i\sigma}=c_{i\sigma}^{\dagger}c_{i\sigma}$ is their corresponding number density, t is the hopping parameter between neighboring sites (denoted by $\langle i,j\rangle$) and U implements the local Coulomb repulsion. The full lattice model was approximated by a two by two cluster embedded in a mean field using the dynamical cluster approximation [17–19]. Using a weak-coupling expansion in continuous imaginary time [20,21] the single-particle Green's function

$$G(i\omega_n)=-\int_0^{\beta}d\tau\, e^{i\omega_n\tau}\langle\mathcal{T}c_i(\tau)c_i^{\dagger}\rangle \tag{45}$$

was calculated for a certain number of Matsubara frequencies $\omega_n=(2n+1)\pi/\beta$. Here $\mathcal{T}$ is the imaginary-time ordering operator, $\langle\cdot\rangle$ denotes a thermal expectation value and $c_i(\tau)=e^{-H\tau}c_ie^{H\tau}$. The model was simulated for $U=W$, where $W=8t$ denotes the bandwidth, and a fixed filling $\langle n_i\rangle=0.9$ for several temperatures T. Within the weak-coupling expansion it is possible to calculate the Green's function directly in frequency space [21] so that no Fourier transformation or discretization of the imaginary-time axis is necessary. In all simulations the number of measured Matsubara frequencies was restricted to $n_{\max}=2U\beta$, which has proven to be sufficient for all calculation. A further increase in the number of frequencies had no influence on the analytic continuation results.

B. Monte Carlo results

Figure 1 shows the α dependence of the single-particle spectra calculated by the parallel tempering Monte Carlo simulation ($\beta=14W^{-1}$). A Gaussian default model

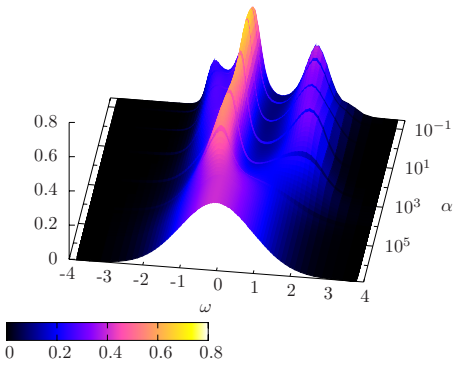

FIG. 1. (Color online) Simulated spectra for a range of regularization parameters α ($\beta=14W^{-1}$). For large α the Gaussian shape of the default model is visible. For decreasing α several features begin to appear.

$$D(\omega)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\omega^2/2\sigma} \tag{46}$$

with $\sigma=1$ was used. The shape of the default model is clearly visible for large α. One can see how several different peaks and other structures appear for decreasing α. Since the α dependence is so strong one definitely needs a criterion to eliminate the regularization parameter.

The density of states calculated by the Wang-Landau simulation and the probability distribution $P[\alpha|\bar{G}]$ following Eq. (39) is shown in Fig. 2. $P[\alpha|\bar{G}]$ is plotted for the two most common choices for $P[\alpha]$, i.e., $P[\alpha]$=constant and $P[\alpha]=1/\alpha$. The density of states varies over at least 15 orders of magnitude (note the logarithmic scales). The probability distributions $P[\alpha|\bar{G}]$ exhibit a well-defined peak at $\hat{\alpha}\sim 0.2$. Note that the two different choices for $P[\alpha]$ have only weak influence on the position of the peak. The two different probability distributions are used to calculate the final single-particle spectrum. Following the discussion in Sec. II, Fig. 2 shows the average of all spectra of Fig. 1 weighted by $P[\alpha|\bar{G}]$ and the spectrum whose α maximizes $P[\alpha|\bar{G}]$. The resulting spectra are nearly indistinguishable and show that neither the ambiguity in the treatment of the probability distribution nor the choice of $P[\alpha]$ has a significant influence on the resulting spectrum.

Let us compare our results from the stochastic analytical inference with those obtained with other methods to fix α. Figure 3 shows that the point where $\chi^2\sim N$ corresponds to $\alpha\sim 1$. That means that the α determined by this rule is identical to Syljuåsen's choice. The χ^2 estimate also exhibits a kink in the very same α region. Thus, the spectra determined by all three methods are identical [Fig. 3(c)]. The chosen $\alpha\sim 1$ is larger than $\hat{\alpha}\sim 0.2$. That indicates that the spectra determined with this criterion are stronger regularized than the spectra calculated by the probability distributions in Fig. 2. However, at least for the QMC data under consideration, the difference between the two spectra is only small. The entropy [Fig. 3(b)] shows no significant features and gives no indication how to choose the α parameter. A sharp drop in the entropy curve is not visible in the simulated area.

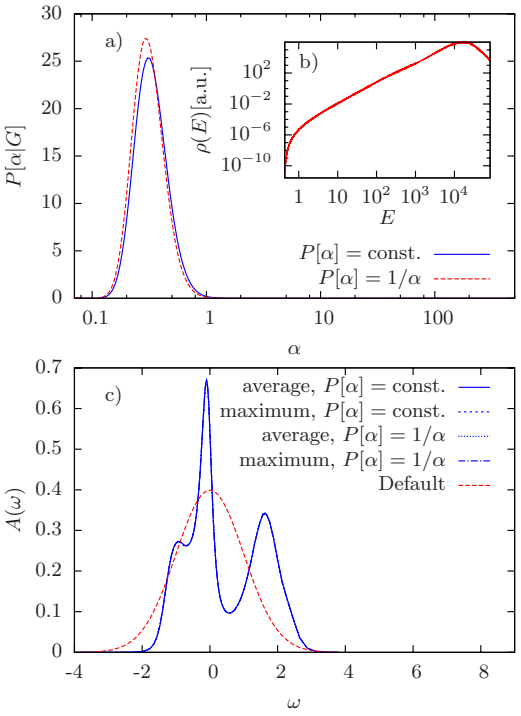

FIG. 2. (Color online) The probability distributions $P[\alpha|\bar{G}]$ (a) based on a Wang-Landau simulation of the density of states (b). The different choices for $P[\alpha]$ only have a weak influence on the position of the peak. The resulting spectra (c) are calculated by either averaging all spectra over $P[\alpha|\bar{G}]$ or by taking the spectrum that maximizes them. The four different spectra are practically identical.

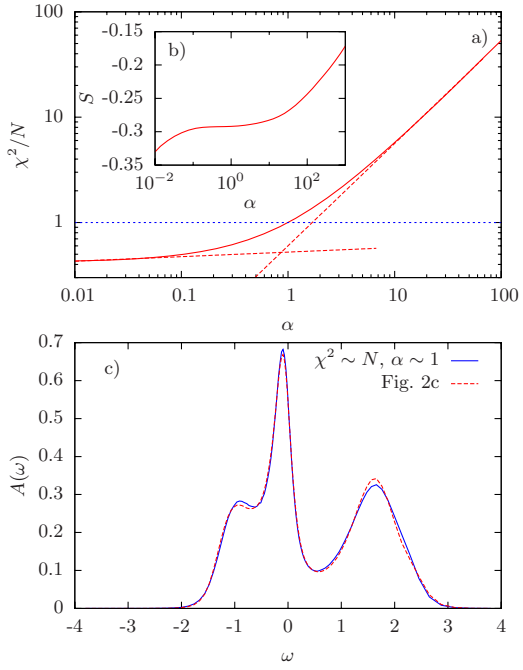

FIG. 3. (Color online) The double-logarithmic plot of χ^2 (a) shows a kink at $\alpha \sim 0.8$ which is very close to the choice of Syljuåsen ($\alpha=1$). It is also the region where $\chi^2 \sim N$. Thus, all three methods give about the same answer and the resulting spectra (c) are identical. A comparison shows that this solution is quite close to the spectrum shown in Fig. 2(c). The entropy (b) exhibits no significant features and gives at least for this data set no indication how to determine α.

Finally, we compare the SAI with the standard MEM approach. Figure 4 shows results of a maximum-entropy calculation using Bryan's algorithm [11] for the same QMC data as before. The qualitative behavior is similar to the SAI simulation: The probability distribution $P[\alpha|\bar{G}]$ shows a noticeable dependence on the prior probability $P[\alpha]$. However, the resulting spectra depend neither on $P[\alpha]$ nor on whether one averages over $P[\alpha|\bar{G}]$ or whether one takes the maximum. The $\chi^2 \sim N$ rule determines an α which is again larger than the one calculated by Bayesian inference. Accordingly the spectrum calculated by this criterion is more regularized, although here the difference is relatively small. Interestingly, in MEM the interesting values for α are about one or two orders of magnitude larger compared to those appearing in the SAI simulations. There seems to be no direct correspondence between the α values of the two methods.

An extended comparison of SAI spectra with results of maximum entropy calculations for several temperatures is collected in Fig. 5. All calculations are based on the Gaussian default model Eq. (46). As already noted before, MEM tends to stronger regularize the spectra and consequently do the SAI spectra exhibit noticeably sharper features for all temperatures shown. Especially the pseudogap, which opens at $\beta=34W^{-1}$, is captured nicely by SAI while the MEM cannot resolve it yet at that temperature.

An important question concerns the dependence of the spectra on the default model. To this end we show in Fig. 6 again SAI and the MEM results for the spectrum at $\beta=34W^{-1}$, this time, however, based on a different default model, namely, a rectangular default model of width 3.6. The resulting spectra are very similar to the one obtained for the Gaussian default model presented in Fig. 5. Thus, even at low temperatures the resulting spectra are quite independent of the default model. More precisely, we could not detect a significant default model dependence at any temperature.

Finally, in order to make a definite statement about the accuracy of our method, we test it on a case where the actual spectrum is known. To this end we create artificial input data by constructing a spectrum

$$A(\omega) = \begin{cases} \dfrac{|x|}{\sqrt{x^2 - \frac{1}{4}}} - 1 & \text{if } |x| > \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}. \tag{47}$$

This function is particularly difficult example for any analytic continuation method, since the actually divergent peak is almost impossible to resolve. The spectrum was transformed into fermionic Matsubara frequency space at $\beta=10$ via the integration kernel in Eq. (4). The first 20 frequencies were calculated. Artificial Gaussian distributed random numbers with a standard deviation $\sigma=5\times10^{-4}$ where added, simulating a somewhat simplified version of statistical errors of real Monte Carlo data by omitting any correlations between the data points. Figure 7 shows that the SAI is more successful in reconstructing the peak than the MEM. It also indicates, that the Wang-Landau criterion developed in

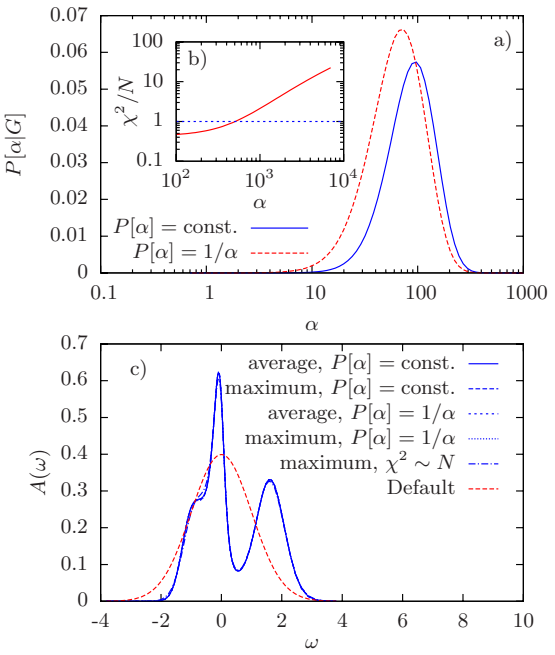

FIG. 4. (Color online) Results of a maximum-entropy calculation for the same QMC data as in Figs. 2 and 3. The probability distribution $P[\alpha|\bar{G}]$ (a) shows a noticeable dependence on $P[\alpha]$, but analogous to SAI, the resulting spectra (c) are identical. The α where $\chi^2 \sim N$ (b) is larger ($\alpha \sim 500$) than the one for which $P[\alpha|\bar{G}]$ is maximal ($\alpha \sim 90$). Accordingly the spectrum chosen by the $\chi^2 \sim N$ rule is more regularized than the one calculated by Bayesian inference.

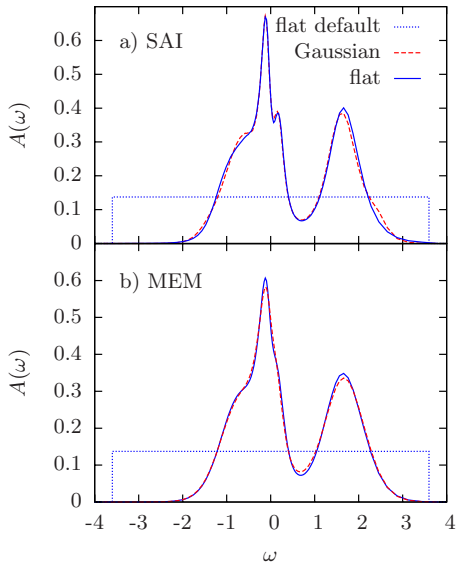

FIG. 6. (Color online) Spectra for β=34W^{-1} based on a flat default model calculated by SAI (a) and by the MEM (b). These spectra are only marginally different compared to those calculated using a Gaussian default model. We conclude that for the QMC data under consideration the calculated spectra are quite independent of the default model.

this paper leads to an even sharper contour and a better reconstruction of the input spectrum compared to the simple $\chi^2 \sim N$ rule.

V. CONCLUSION

We have demonstrated that the stochastic analytic continuation method introduced by Sandvik and Beach can be interpreted in terms of Bayesian probability theory. We developed an algorithm that uses Monte Carlo techniques to both calculate the average spectrum and to eliminate the regularization parameter. It treats all probabilities exactly and hereby avoids the approximations made in the maximum entropy method.

Comparisons to the standard MEM show that the SAI results in robust spectral functions which are less regularized and consequently show more pronounced features, in particular with decreasing temperature in the model calculations. As known from standard MEM, no significant dependence on the default model could be observed. Comparisons to other approaches to fix the regularization parameter α show that the method identifies a smaller α and thus a typi-

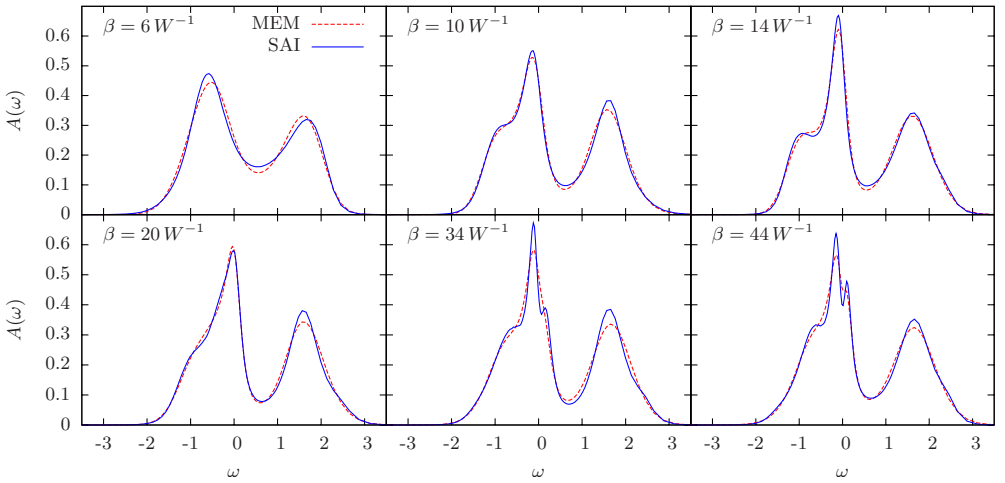

FIG. 5. (Color online) Spectra simulated by stochastic analytical inference compared to maximum-entropy calculations. All calculations are based on a Gaussian default model

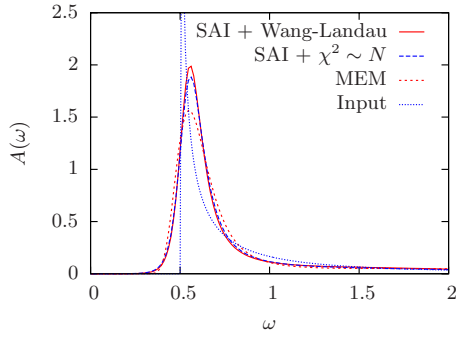

FIG. 7. (Color online) Results for a square-root-shaped spectrum. Only positive frequencies are shown since the spectrum is symmetric around zero. This symmetry was enforced for all simulations. The SAI including the Wang-Landau criterion for the determination of α is the most successful in resolving the sharply peaked input spectrum. Note that the choice $\alpha=1$ gives the same result as $\chi^2 \sim N$.

cally less regularized spectrum, although for the high-quality Monte Carlo data used here the results differ only slightly. Note that this observation adds additional confidence to the method and this type of analytical continuation in general because it proves that in the limit of infinitely precise data, all methods give the same result.

A comparison of the Wang-Landau criterion and the $\chi^2 \sim N$ rule using artificially constructed input data based on a sharply peaked spectrum with a hard gap shows a significant improvement due to the new method; while at the same time the stochastic analytical inference in both cases seems to be more accurate in approximating the singular structure than the classical maximum entropy. As to when simple rules such as $\chi^2 \sim N$ or the choice $\alpha=1$ may or may not present a good way to fix the regularization parameter depends sensitively on the model, the quality of the data, and structures occurring in the spectral function. However, we can expect the stochastic analytical continuation to result in spectra which are in general closer to the exact one, with the Wang-Landau approach typically giving the most accurate image.

One apparent drawback of the method is the necessity to perform simulations for a broad range of values for α, independent of whether one chooses the Wang-Landau approach or $\chi^2 \sim N$, respectively, $\alpha=1$ [23] to fix α. Although this can be performed efficiently with parallel tempering techniques, the required computer resources for one single spectrum can sum up to about 20 processor hours and are hence orders of magnitude larger than for standard MEM approaches. Especially for QMC data at higher temperatures, more computer time may be needed for the analytic continuation than for the simulation of the Monte Carlo data itself. As the resulting spectra tend to be less regularized one has to ponder the gain in details in the structures against the significant increase in computer time.

ACKNOWLEDGMENTS

Our implementation of all algorithms is based on the libraries of the ALPS project [22]. ALPS (Applications and Libraries for Physics Simulations, http://alps.comp-phys.org) is an open source effort providing libraries and simulation codes for strongly correlated quantum-mechanical systems. We acknowledge financial support by the Deutsche Forschungsgemeinschaft through Grant No. SFB 602 and by the German Academic Exchange Service (DAAD). This work was supported by the National Science Foundation through Grants No. OISE-0952300 and No. DMR-0706379. Portions of this research used computational resources provided by the Louisiana Optical Network Initiative (http://www.loni.org) and by the Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (http://www.gwdg.de).

[1] M. Jarrell and J. Gubernatis, Phys. Rep. **269**, 133 (1996).
[2] E. T. Jaynes, in *Maximum-Entropy and Bayesian Methods in Applied Statistics*, edited by J. H. Justice (Cambridge University Press, Cambridge, 1986), p. 1.
[3] S. F. Gull, in *Maximum-Entropy and Bayesian Methods in Science and Engineering*, edited by G. J. Erickson and C. R. Smith (Kluwer Academic Press, Dordrecht, 1988), Vol. 1, p. 53.
[4] A. W. Sandvik, Phys. Rev. B **57**, 10287 (1998).
[5] K. Beach, e-print arXiv:cond-mat/0403055.
[6] O. F. Syljuåsen, Phys. Rev. B **78**, 174429 (2008).
[7] A. Papoulis, *Probability and Statistics* (Prentice Hall, New York, 1990), p. 422.
[8] J. Skilling, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer Academic Press, Dordrecht, 1989), p. 45.
[9] J. Skilling, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer Academic Press, Dordrecht, 1989), p. 53.
[10] J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, Phys. Rev. B **44**, 6011 (1991).
[11] R. K. Bryan, Eur. Biophys. J. **18**, 165 (1990).
[12] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[13] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Vel'yaminov, J. Chem. Phys. **96**, 1776 (1992).
[14] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[15] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[16] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[17] M. H. Hettler, A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998).
[18] M. H. Hettler, M. Mukherjee, M. Jarrell, and H. R. Krishnamurthy, Phys. Rev. B **61**, 12739 (2000).
[19] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005).
[20] A. N. Rubtsov and A. I. Lichtenstein, JETP Lett. **80**, 61 (2004).
[21] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Phys. Rev. B **72**, 035122 (2005).
[22] A. F. Albuquerque *et al.*, J. Magn. Magn. Mater. **310**, 1187 (2007).
[23] This is mandatory to ensure good convergence of the Monte Carlo process.