Application of Bayesian Inference to Stochastic Analytic Continuation

S Fuchs¹, M Jarrell² and T Pruschke¹

¹ Institut für Theoretische Physik, Georg-August-Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany, ² Center of Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

E-mail: fuchs@theorie.physik.uni-goettingen.de

Abstract. We present an algorithm for the analytic continuation of imaginary-time quantum Monte Carlo data. The algorithm is strictly based on principles of Bayesian statistical inference. It utilizes Monte Carlo simulations to calculate a weighted average of possible energy spectra. 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.

1. Introduction

Quantum Monte Carlo simulations are powerful tools to calculate properties of strongly correlated electron systems. Of particular interest in those systems are dynamical correlation functions like single-particle spectra. However, QMC presently provides data only on the imaginary time axis, and the necessary analytic continuation of these data has proven to be difficult. Presently, the Maximum Entropy Method (MEM) is the standard tool to solve this problem [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. We show that this stochastic approach can also be understood in terms of Bayesian statistical inference and provide a strict criterion completely based on Bayesian logic to eliminate the free parameter. We apply the algorithm to imaginary-time quantum Monte Carlo data and compare the resulting energy spectrum with a standard maximum entropy calculation.

2. The Maximum Entropy Method

Quantum Monte Carlo simulations can provide accurate estimates \( \tilde{G}_i = \tilde{G}(\tau_i) \) for an imaginary-time correlation function \( G(\tau) \) at a finite set of \( N \) imaginary-time points \( \tau_i \). The task of calculating the real-time correlation function \( G(t) \) from this information is called analytic continuation. The input to the analytic continuation procedure consists of these Monte Carlo estimates \( \tilde{G}_i \) and their covariance matrix \( C_{ij} = \tilde{G}_i \tilde{G}_j - \tilde{G}_i \tilde{G}_j \). In principle the spectral function
\[ A(\omega) = -\frac{1}{\pi} \text{Im} \ G(\omega + i0^+) \] can be extracted from these data by inverting

\[ G(\tau) = \int d\omega K(\tau, \omega) A(\omega) \quad \text{with} \quad K(\tau, \omega) := -\frac{e^{-\omega \tau}}{1 \pm e^{-\omega/(kT)}}, \tag{1} \]

where the upper sign holds for fermions and the lower one for bosons. A least-square fit minimizes

\[ \chi^2[A] = \frac{1}{N} \sum_{i,j} \left( G_i - G(\tau_i) \right) \sqrt{C_{ij}^{-1}} \left( \hat{G}_j - G(\tau_j) \right) \]

with respect to \( A(\omega) \). The maximum entropy method can be understood as an attempt to regularize this least-square fit. One defines an entropy \( S[A] = -\int d\omega A(\omega) \ln A(\omega)/D(\omega) \) relative to a default model \( D(\omega) \). Instead of just minimizing \( \chi^2 \) the MEM minimizes the quantity \( Q[A] = \chi^2[A] - \alpha S[A] \), introducing a regularization parameter \( \alpha \). The MEM can be reformulated using Bayesian probability theory [7] by defining subjective probabilities for the quantities involved [2, 3]. This way the probability of the spectrum \( A \) given the input data can be shown to be \( P[A|G] \sim e^{-Q[A]} \). Thus the MEM determines the most probable spectrum \( \hat{A}_\alpha \), given \( G \). Bayesian inference can also be used to find an expression for \( P[\alpha|\hat{G}] \), the probability of \( \alpha \) given \( \hat{G} \), which provides criteria to eliminate the parameter \( \alpha \) [8, 9].

3. Stochastic Analytical Inference

An alternative to the MEM [4, 6] was proposed, that does not employ the explicit regularization of the fit by an entropy term. Rather than maximizing \( P[A|\hat{G}] \), an average over all possible spectra weighted by \( \exp(-\chi^2/\alpha) \) is performed. Beach refined this approach, by introducing the default model \( D(\omega) \) of the MEM into the algorithm [5]. By defining a field \( n = A/D \) and calculating the average field

\[ \langle n \rangle_\alpha = \frac{1}{Z} \int \mathcal{D}n ne^{-\chi^2[n]/\alpha} \quad \text{with} \quad Z = \int \mathcal{D}n e^{-\chi^2[n]/\alpha}, \tag{3} \]

the average spectrum \( \langle A \rangle_\alpha \) can be regained from \( \langle A \rangle_\alpha = D\langle n \rangle_\alpha \). Similar to the MEM [1, 8, 9] Bayesian logic can be utilized to calculate the probability \( P[\alpha|\hat{G}] \). Bayes’s Theorem [7] connects the relevant probabilities for the field \( n \) and the regularization parameter \( \alpha \):

\[ P[n, \alpha|\hat{G}] = P[\hat{G}|n, \alpha] \ P[n, \alpha] / P[\hat{G}] \tag{4} \]

Analogous to the MEM one identifies \( P[\hat{G}|n, \alpha] = e^{-\chi^2/n}/Z' \) with \( Z' = \int d\hat{G} e^{-\chi^2/\alpha} \sim \alpha^{N/2} \). \( P[\hat{G}] \) is an \( \alpha \)-independent normalization constant. \( P[n|\alpha] \), which is normally associated with the entropy, is taken to be constant, since there is no explicit entropy regularization in the stochastic approach under consideration. One obtains

\[ P[\alpha|\hat{G}] = P[\alpha] \int \mathcal{D}n P[\hat{G}|n, \alpha] \ P[n|\alpha] / P[\hat{G}] \sim P[\alpha] \alpha^{-N/2} \int \mathcal{D}n e^{-\chi^2[n]/\alpha}. \tag{5} \]

All quantities in this equation are known except \( P[\alpha] \). Analogous to the MEM it is either taken to be constant or to be the Jeffreys prior \( 1/\alpha \) [8–10].

4. Simulation Results

The average (3) is calculated by a standard Monte Carlo simulation using Metropolis weights. The simulation is performed for a wide range of \( \alpha \)-values using a parallel tempering algorithm [11] to ensure convergence for small \( \alpha \). A numerical treatment of (5) involves the calculation of
Figure 1. Simulated spectra for a range of regularization parameters $\alpha$. For large $\alpha$ the Gaussian shape of the default model is clearly visible. For decreasing $\alpha$ several features begin to appear.

$Z = \int Dn e^{-\chi^2/\alpha}$. This is equivalent to calculating a partition function in a canonical ensemble. We use a Wang-Landau algorithm [12, 13] to simulate the density of states $\rho(E)$ of the system. Afterwards the partition function can be obtained by $Z = \int dE \rho(E) e^{-E/\alpha}$. Our implementation follows [5] closely and is based on the libraries of the ALPS project [14].

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

$$H = -t \sum_{\langle i,j \rangle} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (6)$$

here $i$ and $j$ are lattice site indices, the operators $c_{i\sigma}^\dagger$ ($c_{i\sigma}$) create (destroy) an electron with spin $\sigma \in \{\uparrow, \downarrow\}$ on site $i$. $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ is their responding number density, $t$ is hopping parameter between neighbouring sites (denoted by $\langle i, j \rangle$) and $U$ is 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 [15, 16]. The single-particle Green function $G(\tau) = -\langle T c_i(\tau) c_i^\dagger \rangle$ of the cluster model was calculated using a Hirsch-Fye quantum Monte Carlo algorithm [17–19]. Here $T$ is the imaginary-time ordering operator, $\langle \cdot \rangle$ denotes a thermal expectation value and $c_i(\tau) = e^{-H\tau} c_i e^{H\tau}$. The model was simulated for $U = 8t$ and a fixed filling $\langle n \rangle = 0.8$ for an inverse temperatures $\beta = 1/k_B T = 7t^{-1}$.

Fig 1 shows the $\alpha$-dependence of the single particle spectra calculated by the parallel tempering Monte Carlo simulation using a default model of a Gaussian shape. The density of states calculated by the Wang-Landau simulation and the probability distribution $P[\alpha|\bar{G}]$ following (5) is shown in Fig. 2. $P[\alpha|\bar{G}]$ is plotted for the two most common choices for $P[\alpha]$, i.e. $P[\alpha] = \text{const}$ and $P[\alpha] = 1/\alpha$. One has always two possibilities to treat $P[\alpha|\bar{G}]$: (i) One calculates $\alpha^*$ as the $\alpha$ that maximizes $P[\alpha|\bar{G}]$ and takes $\langle n \rangle_{\alpha^*}$ as the final result. (ii) One averages over all $\langle n \rangle_\alpha$ weighted by the probability $P[\alpha|\bar{G}]$. Fig. 3 shows that neither this ambiguity nor the choice of $P[\alpha]$ have a noteworthy influence on the resulting spectrum.

5. Conclusion

We show that the stochastic analytic continuation method introduced by Sandvik and Beach can be interpreted in terms of Bayesian probability theory. We develop 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. Fig. 3 shows that the method is able to produce spectra
Figure 2. The density of states calculated by a Wang-Landau simulation (a) and the probability distribution $P[\alpha|\bar{G}]$ (b).

Figure 3. The spectrum simulated by Stochastic Analytical Inference, calculated by either averaging all spectra over the probability distributions $P[\alpha|\bar{G}]$ shown in Fig. 2 or by taking the spectrum that maximizes them. A result of a MEM calculation using Bryan’s algorithm [9] is also shown.

with very similar but slightly sharper features compared to the result of a maximum entropy calculation.

Acknowledgments
We acknowledge financial support by the Deutsche Forschungsgemeinschaft through SFB 602 and by the Deutscher Akademischer Austausch Dienst (DAAD).

References
[1] Jarrell M and Gubernatis J 1996 Phys. Rep. 269 133
[2] Jaynes E T 1986 Maximum Entropy and Bayesian Methods in Applied Statistics ed Justice J H (Cambridge: Cambridge University Pr.) p 1
[3] Gull S F 1988 Maximum Entropy and Bayesian Methods in Science and Engineering vol 1 ed Erickson G J and Smith C R (Dordrecht: Kluwer Acad. Pr.) p 53
[4] Sandvik A W 1998 Phys. Rev. B 57 10287
[5] Beach K 2004 Cond-mat/0403055 (unpublished)
[6] Syljuåsen O F 2008 Phys. Rev. B 78 174429
[7] Papoulis A 1990 Probability and Statistics (New York: Prentice Hall) p 422
[8] Skilling J 1989 Max. Ent. and Bayesian Methods ed Skilling J (Dordrecht: Kluwer Acad. Pr.) pp 45,51
[9] Bryan R K 1990 Eur. Biophys. J. 18 165
[10] Gubernatis J E, Jarrell M, Silver R N and Sivia D S 1991 Phys. Rev. B 44 6011
[11] Swendsen R H and Wang J 1986 Phys. Rev. Lett. 57 2607
[12] Wang F and Landau D P 2001 Phys. Rev. Lett. 86 2050
[13] Wang F and Landau D P 2001 Phys. Rev. E 64 056101
[14] Albuquerque A F and et al for the ALPS Collaboration 2007 JMMM 310 1187–1193
[15] Hettler M, Tahvildar-Zadeh A, Jarrell M, Pruschke T and Krishnamurthy H 1998 Phys. Rev. B 58 R7475
[16] Maier T, Jarrell M, Pruschke T and Hettler M H 2005 Rev. Mod. Phys. 77 1027
[17] Hirsch J and Fye R 1986 Phys. Rev. Lett. 56 2521
[18] Jarrell M 1992 Phys. Rev. Lett. 69 168
[19] Jarrell M, Maier T, Huscroft C and Moukouri S 2001 Phys. Rev. B 64 195130