

Analytical continuation of imaginary axis data using maximum entropy

O. Gunnarsson, M.W. Haverkort and G. Sangiovanni

For strongly correlated systems analytical methods usually involve uncontrolled approximations. Therefore stochastic methods such as quantum Monte-Carlo (QMC) or quantum cluster methods are often used. Apart from statistical errors, these methods often produce very accurate results, but the results are obtained on the imaginary time or frequency axis. This leaves the problem of analytically continuing the results to the real frequency axis, which is an ill-posed problem. Small changes in the data on the imaginary axis can lead to large changes on the real axis. Since the imaginary axis data contain statistical noise, the analytical continuation is very difficult.

This problem can be treated within the Bayesian theory [1]. The problem is regularized by introducing an entropy in terms of the deviation of the real axis output spectrum from some default function. The importance of the entropy is controlled by a parameter α , which is determined using statistical arguments [1]. This method is referred to as the Maximum Entropy (MaxEnt) method. It has been rather successful in performing analytical continuations.

We have studied [2] how the accuracy of this method can be improved. The error in the output spectral function can be split up in a statistical error, due to the noise in the input data, and a systematic error, due to the deviation of the default function from the true spectrum. The choice of α determines the relative size of these errors. In the classical MaxEnt method the most probable α is chosen [1]. We find that this choice can make the statistical error unnecessary large.

The input data is typically given as a number N_{sample} of samples, $\overline{G}_V(\tau)$, where each sample gives a (noisy) version of the imaginary time function $G(\tau)$. We find that the accuracy can sometimes be improved by splitting up the samples in N_{calc} sub sets (batches). We then perform N_{calc} MaxEnt calculations, each with

$N_{\text{sample}}/N_{\text{calc}}$ samples, and then average the results, instead of performing one MaxEnt calculation with N_{sample} samples. This ‘batching’ approach reduces the statistical error at the cost of an increase in the systematic error.

We have also studied the possibility of an iterative MaxEnt method, where the output is used to define a new default function. This usually works poorly. Although there is an improvement in the systematic error, we find that this is usually overwhelmed by an increase in the statistical error. However, if the data are split in batches, as discussed above, the importance of the statistical error can be reduced to the point where the approach improves the total accuracy.

We focus on a response function, the optical conductivity $\sigma(\omega)$. A typical $\sigma(\omega)$ is introduced, which in the following will be referred to as the ‘exact’ $\sigma(\omega)$. This model of $\sigma(\omega)$ can be transformed accurately to the imaginary axis. Statistical noise with the magnitude σ_0 are added to the imaginary axis data, which is then transformed back to the real axis, using various modifications of the MaxEnt method. If a given method worked perfectly, the $\sigma(\omega)$ that we started from would be recovered. The deviation

$$w = \int_0^\infty [\sigma_{\text{exact}}(\omega) - \sigma_{\text{calc}}(\omega)]^2 d\omega \quad (1)$$
$$\equiv w_{\text{stat}} + w_{\text{sys}},$$

from the ‘exact’ $\sigma(\omega)$ is then a measure of the accuracy of different approaches. Here we have also introduced statistical and systematic errors.

The inset of Fig. 1 shows the default model used to calculate the entropy as well as the the model optical conductivity. From this model, we generate data for imaginary time with 100 samples, each with the statistical noise σ_0 . For $\sigma_0 = 0.01$, a classical MaxEnt calculation prescribes the value $\alpha = 40$, which is used in the following.

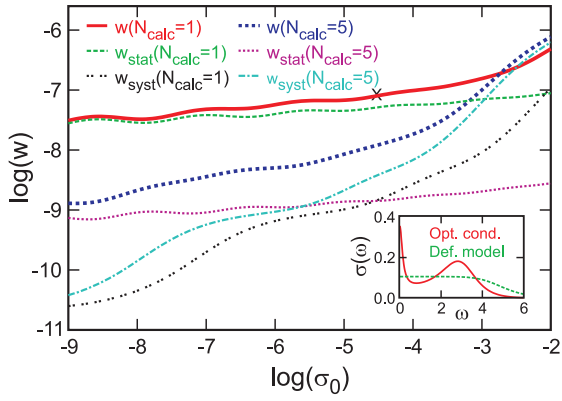


Figure 1: Statistical (w_{stat}), systematic (w_{syst}) and total (w) errors in MaxEnt calculations for the spectrum and default model in the inset, using 100 samples, each with the accuracy σ_0 . The thick full red line shows w when one MaxEnt calculation is performed using all 100 samples and the thick broken blue line the result when 5 MaxEnt calculations are averaged, each calculation using 20 samples. The cross corresponds to a classical MaxEnt calculation for $\sigma_0 = 0.01$, which gives $\alpha \approx 40$, used in the figure. The figure illustrates that the error can be substantially reduced by batching the data.

Figure 1 shows the statistical and systematic errors as a function of σ_0 . The figure illustrates that the statistical error (green dotted curve) is much larger than the systematic error (black dotted curve). This is also illustrated in Fig. 2(a). This shows the result of 20 MaxEnt calculations with different realizations of the noise, each with 100 samples with the accuracy $\sigma_0 = 0.01$. The thick red line shows the exact spectrum. The calculated spectra (thin green lines) scatter strongly around the exact result, illustrating the large statistical error. On the average, these spectra also deviate somewhat from the exact result, illustrating a small systematic error.

We next group the 100 samples in $N_{\text{calc}} = 5$ batches, each with 20 samples, and perform 5 MaxEnt calculations. The statistical noise of the data in each MaxEnt calculations is then increased by a factor $\sqrt{5}$. This increases both the

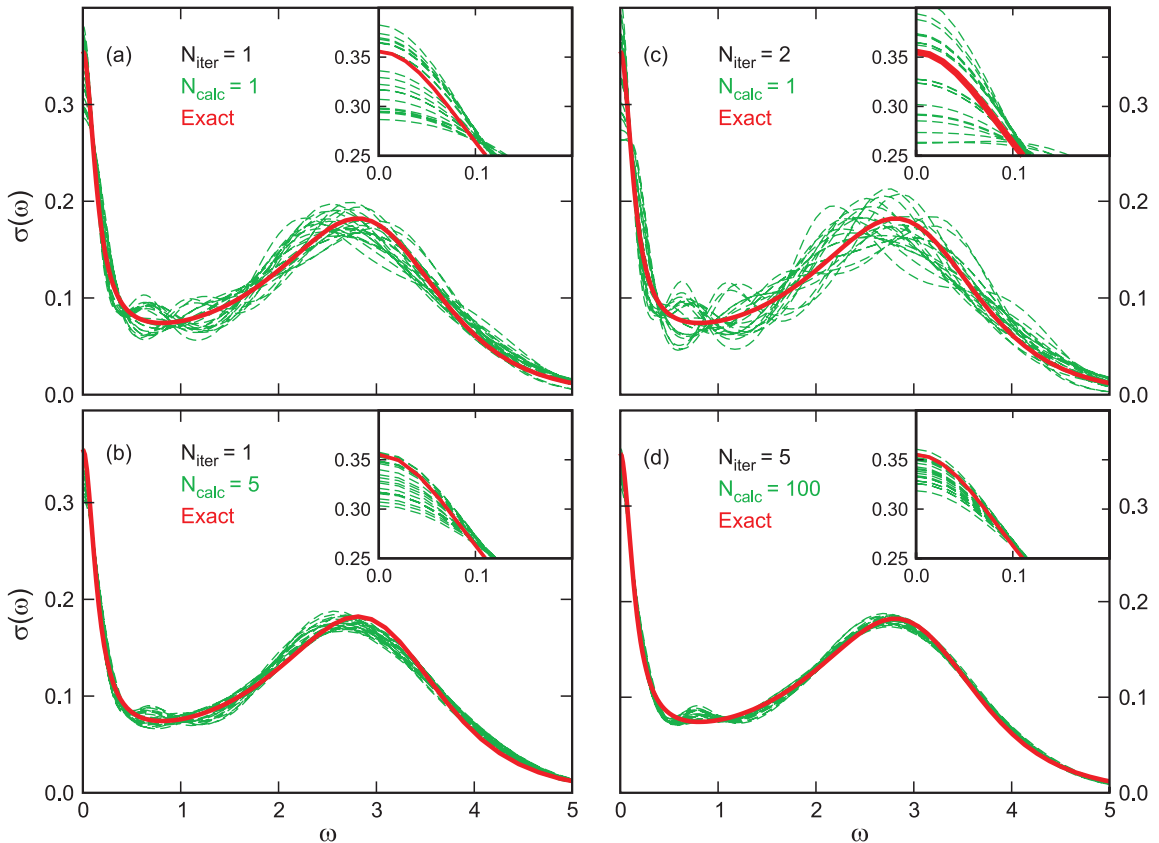


Figure 2: Optical conductivity using different methods. 100 samples, each with the accuracy $\sigma_0 = 0.01$ were given. (a) Each curve shows results of a classical MaxEnt calculation using all 100 samples. The figure shows 20 such curves, each corresponding to a different realization of the noise. (b) Each curve shows the average of five MaxEnt calculations, each using 20 samples. (c) Each curve shows the results of iterating the calculations in (a) once, using the output in (a) as a default function in the next MaxEnt calculation. (d) Each curve shows the results of iterating MaxEnt calculations five times. $N_{\text{calc}} = 100$ was used, and the default function was obtained from the average of these N_{calc} calculations.

systematic and statistical errors of each calculation somewhat. Averaging these calculations, however, reduces the statistical error by a factor 5.

In Fig. 1 this leads to a large net reduction in the statistical error (dotted red line), which more than compensates for the increase of the systematic error (thin chain blue line). This is also illustrated in Fig. 2(b), which shows 20 such results, each one obtained by averaging 5 MaxEnt calculations with 20 samples, but with different realizations of the noise. As is also illustrated in Tab. 1, the spread between the curves is substantially smaller than in Fig. 2(a), while the systematic error is somewhat larger. This leads to a substantial improvement in the total error (thick broken blue curve in Fig. 1).

Table 1: Statistical (w_{stat}), systematic (w_{syst}) and total (w) errors in the present MaxEnt calculations. 100 samples, each with the accuracy $\sigma_0=0.01$, were split up in N_{calc} batches with $100/N_{\text{calc}}$ samples and used in N_{calc} calculations. The average of the outputs was used as a default model, performing N_{iter} iterations.

N_{calc}	N_{iter}	$w_{\text{stat}} \times 10^4$	$w_{\text{syst}} \times 10^4$	$w \times 10^4$
1	1	6.7	1.4	8.1
5	1	1.4	2.1	3.5
1	2	14	1.3	15
100	5	0.4	0.9	1.3

To provide a criterion for how to split up the data in batches, we have estimated the change in the statistical error when N_{calc} is changed [2]. We compare calculations with $N_{\text{calc}} = N_{\text{sample}}$ and $N_{\text{calc}} = N_{\text{sample}} / 2$ batches. In the second calculation the statistical error is larger and the systematic error is smaller. If the total difference between the two calculations is larger than twice the estimated statistical difference, this suggests that the reduction of the systematic error outweighs the increase of the statistical error and the second calculation is assumed to be more accurate. In this way N_{calc} is reduced until this prescription predicts no further improvement.

Once a MaxEnt calculation has been performed, one can try to improve the default function by using the output spectral function as a new default function. Such an iterative approach, how-

ever, is usually found to lead to poor results, as is illustrated in Fig. 2(c). The spread between different calculations is larger than in the non-iterated case in Fig. 2(a), due to an increased statistical error. We next split up the samples in N_{calc} batches. Performing $N_{\text{iter}} = 5$ iterations, the total error is reduced, as is shown in Fig. 2(d). The use of $N_{\text{calc}} = 100$ drastically reduces the statistical error. The following iterations increase the statistical error substantially, but it nevertheless remains small. At the same time the iterations reduce the systematic error, so that both are improved compared with the non-iterated case (see Tab. 1 and compare Fig. 2(d) with Figs. 2(a) and (b)). We have not found a reliable criterion for when to stop the iterations.

We are now in the position to discuss the limits of accuracy that can be obtained in this approach. As before, we consider 100 samples with the accuracy $\sigma_0=0.01$ and allow for any combination of α , $N_{\text{calc}} \leq 100$ and $N_{\text{iter}} \leq 40$. Since the exact result is known, we can test which combination gives the smallest error. This leads to the results shown in Fig. 3.

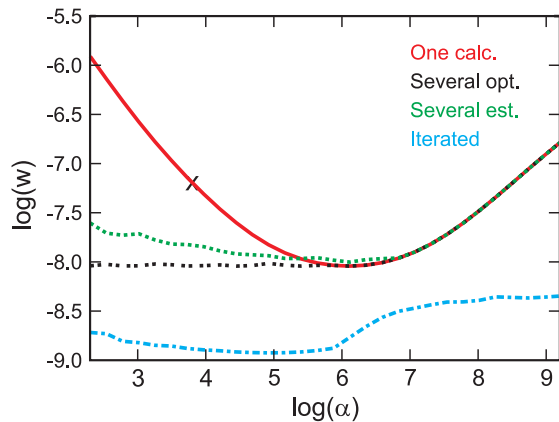


Figure 3: Accuracy w of MaxEnt calculations for 100 samples, each with accuracy $\sigma_0=0.01$. ‘One calc.’ uses all samples in one MaxEnt calculation. ‘Several opt.’ and ‘Several est.’ split up the samples in several batches and average the resulting MaxEnt calculations. ‘Several opt.’ does this in the optimal way and ‘Several est.’ uses a prescription for finding the splitting when the exact result is not known. ‘Iterated’ in addition uses the output spectral function as default model in an iterative approach. The cross shows the result of a classical MaxEnt calculation.

The curve ‘One calc.’ shows the result of a traditional MaxEnt calculation, using all the samples in one calculation. A classical MaxEnt calculation prescribes $\alpha \approx 40$, which is shown by

a cross. This value of α is not optimal, and a larger α would have given a smaller error. We next introduce batching of the samples. We find the number of batches which gives the best agreement with the ‘exact’ $\sigma(\omega)$. This (‘Several opt.’) leads to a much higher accuracy for small values of α . For large values of α , one batch gives the best accuracy, and the curve falls on top of the curve ‘One calc.’. In actual applications, when the exact result is not known, the criterion for dividing the samples in batches has

to be used. This approach (‘Several est.’) leads to almost as good results. Finally, the curve ‘Iterated’ shows results when iterations are allowed. This leads to a substantial improvement in the accuracy.

References:

- [1] *Jarrell, M. and J.E. Gubernatis*. Physics Reports **269**, 133–195 (1996).
- [2] *Gunnarsson, O., M.W. Haverkort and G. Sangiovanni*. Physical Review B **81**, 155107 (2010).