# 1 Maximum entropy method

The numerically exact impurity solvers used in solving quantum many-body problems return results on the imaginary time or imaginary frequency axis. To obtain data on the real frequency axis, one must perform an analytic continuation from the imaginary to the real axis. Several methods exist for this purpose; here, we use the *maximum entropy method*. Here we mainly perform the analytic continuation for the modified periodic Anderson model, in the same phase regime as discussed in papers (Phys. Rev. B 107, 205104 and arXiv:2509.00225). To briefly outline, the model exhibits a continuous quantum phase transition with a critical point at $t_\perp = 1.033$ separating the Fermi liquid metal ($t_\perp < 1.033$) from the Mott insulating phase ($t_\perp > 1.033$). Other parameters relevant to the model are the same as in the paper.

## 1.1 Introduction to the problem of analytic continuation

Understanding quantum interacting systems at finite temperature is of fundamental importance. In condensed matter physics, one of the most widely used numerically exact methods for studying such systems at a finite temperature is the Continuous-Time Quantum Monte-Carlo(CTQMC). CTQMC operates in Matsubara(imaginary) time or frequency domains and provides highly accurate results. However, since most experimental observables are measured in real frequencies, it becomes essential to perform analytic continuation of the CTQMC data to obtain real frequency response functions. While one can sometimes infer qualitative features of real-frequency behaviour directly from the imaginary-frequency domain. In principle, analytic continuation is exact if the Matsubara data is noise-free. However, CTQMC is inherently stochastic, and the resulting data contains noise, particularly at high frequencies and low temperatures. This renders the analytic continuation problem ill-posed, as there is no unique solution.

Several methods have been developed to address this challenge, including Padé approximation [1, 2], Nevanlinna Analytical Continuation[3], singular value decomposition [4], Maximum Entropy Methods(MEM)[5–9], and more recently, machine learning approaches[10–13]. For data with low noise, the Padé approximation often yields excellent results. In this method, the function is approximated by a ratio of two polynomials, with constraints derived from the underlying physical properties. However, in general CTQMC simulations, especially at low temperatures and high frequencies, the data is noisy, and Padé tends to overfit, producing spurious features.

The Maximum Entropy Method (MEM) is a well-established method, developed nearly three decades ago. Since the underlying calculation is inherently stochastic, it is natural to exploit this stochasticity in the analytic continuation process, precisely what MEM does. By incorporating the statistical nature of the data, MEM effectively narrows the space of possible solutions. However, this comes at the cost of requiring accurate statistical information about the data, which is computationally demanding to obtain. Nevertheless, MaxEnt performs remarkably well on the moderate to low precision data. In this chapter, we employ MaxEnt to perform the analytic continuation of the CTQMC data.

The equations relevant to the analytic continuation problem are given by:

$$G(i\omega_n) = \int d\omega \frac{1}{i\omega_n - \omega}\rho(\omega) \tag{1}$$

$$G(\tau) = \int d\omega \frac{e^{-\omega\tau}}{1 + e^{-\omega\beta}}\rho(\omega) \tag{2}$$

Here, $G(i\omega_n)$ or $G(\tau)$ represents the input data, from which the goal is to extract the spectral function $\rho(\omega)$. When both $G(i\omega_n)$ and $G(\tau)$ are available, it is generally preferable to work with $G(i\omega_n)$. This is because the kernel in Eq. 2 leads to integrals involving complicated hypergeometric functions. Furthermore, the kernel in Eq. 2 becomes increasingly sharp around $\omega = 0$ as the temperature decreases, making numerical integration challenging. Therefore, $G(i\omega_n)$ is typically the more suitable choice for analytic continuation [6].

## 1.2 Method

The MEM employs a Bayesian framework to obtain a unique solution to the analytic continuation problem. For a given input dataset $G$ with associated error $\sigma$, prior information can be used to constrain the resulting spectral function $\rho(\omega)$ to the space defined by $G - \sigma < \mathbf{K}\rho < G + \sigma$, following Bayes theorem:

$$P(\rho|G) = \frac{P(G|\rho)P(\rho)}{P(G)} \tag{3}$$

Here, $P(\rho|G)$ is the posterior probability, which is the probability of obtaining the spectral function $\rho$ given the data $G$. $P(G|\rho)$ is the likelihood, representing the probability of observing $G$ for a given $\rho$. The term $P(G)$ is a normalization constant, as the input data is fixed. $P(\rho)$ is the prior probability, incorporating prior knowledge about the spectral function. The goal is to maximize $P(\rho|G)$ to find the most probable $\rho$.

According to the central limit theorem, the likelihood is given by,

$$P(G|\bar{G}) \propto e^{-\chi^2/2}, \qquad \chi^2 = \sum_i^N \frac{(G_i - \bar{G})^2}{\sigma_i^2} \tag{4}$$

where $G_i$ is the $i^{\text{th}}$ sample, $\bar{G}$ is the sample average, and $N$ is the the total number of data points. If the covariance matrix is non-diagonal, one uses the generalized form: $\chi^2 = (G-\bar{G})^T C^{-1}(G-\bar{G})$, where $C$ is the covariance matrix. The prior $P(A)$ should be chosen to encode relevant information such as positivity and normalization (sum rules) of the spectral function. A common choice is $P(\rho) \propto e^{\alpha S}$[14, 15], where $S$ is the relative entropy given by

$$S = -\sum_k \frac{\Delta\omega_k}{2\pi}\rho(\omega_k)\ln\frac{\rho(\omega_k)}{\rho_0(\omega_k)}$$

Here, $\rho_0(\omega)$ is the default model, which encodes prior assumptions about the spectral function. The posterior probability to be maximized thus takes the form

$$P(\rho|G) \propto e^{\alpha S - \chi^2/2} \tag{5}$$

To determine the optimal value of $\alpha$, several approaches exist, such as the classic method [16] and Bryan's approach[8]. Both require the default model to be close to the true spectral function. Here we adopt an alternative method proposed in Ref. [6] to find the optimum value of the $\alpha$, which doesn't rely on the default model being close to the actual spectral function.

Examining Eq.5, one can see a competition between two terms: the $\chi^2$ term, which enforces consistency with the input data, and the relative entropy term $S$, which encodes proximity to the default model. In the limiting case $\alpha = 0$, the spectrum is entirely determined by the data, including noise. Hence, this is the noise fitting regime. Conversely, for $\alpha = 1$, the data are ignored, and the result reflects the default model. This is the entropy-dominated regime (see Fig. 1). Between these extremes lies an intermediate regime where $\log\chi^2$ varies approximately linearly with $\log\alpha$; this is termed the information fitting region. The optimal $\alpha$ corresponds to the point where the transition from information-fitting to noise-fitting occurs. At this point, the method identifies and suppresses noise, thereby fitting only the true data. The optimal value of $\alpha$ can be determined by fitting straight lines to the $\log\chi^2$ vs. $\log\alpha$ curves in the information-fitting and Noise-fitting regions and extracting their intersection[6].

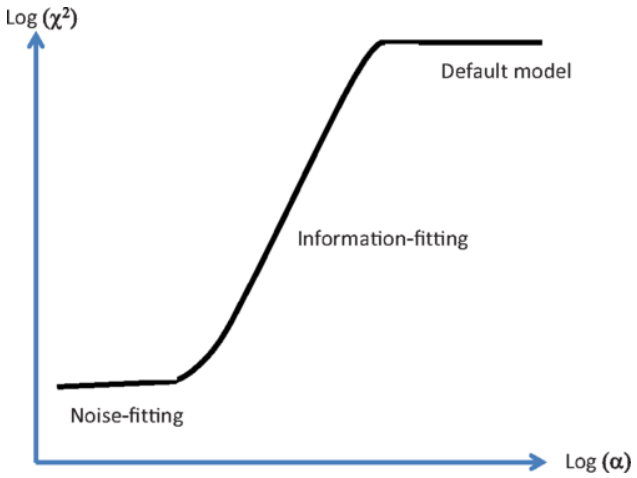


Figure 1: Schematic plot of $\log(\chi^2)$ versus $\log(\alpha)$ in the Maximum Entropy Method. The optimal value of $\alpha$ is chosen at the 'kink' separating the information-fitting regime, where the fit is sensitive to the signal in the data, from the noise-fitting regime, where the fit begins to incorporate statistical noise.

## 1.3 Preparation of the input data

The first step is to obtain well-converged data, which is achieved by performing finite temperature calculations for a modified periodic Anderson model within the framework of DMFT. These calculations are carried out over many iterations with a large number of QMC cycles to ensure convergence. Once convergence is reached, we estimate the error statistics required for the MaxEnt procedure as follows.

We generate 1000-2000 input files, each initialized with a different random seed. Using different seeds is essential to obtain uncorrelated data. The converged result is used as the starting point for all these files. Each output from this run is treated as a statistical bin. Random seeds are applied not only across input files but also across different processors in parallel runs, which helps eliminate correlations among the data sets. To verify the statistical quality of the resulting data, we assess the third and fourth-order moments, namely skewness and relative kurtosis. Skewness quantifies the asymmetry of the distribution (with zero indicating perfect symmetry), while relative kurtosis measures the tailedness of the distribution relative to that of a Gaussian distribution. In our analysis, both skewness and relative kurtosis are found to be close to zero, confirming that the data are normally distributed.

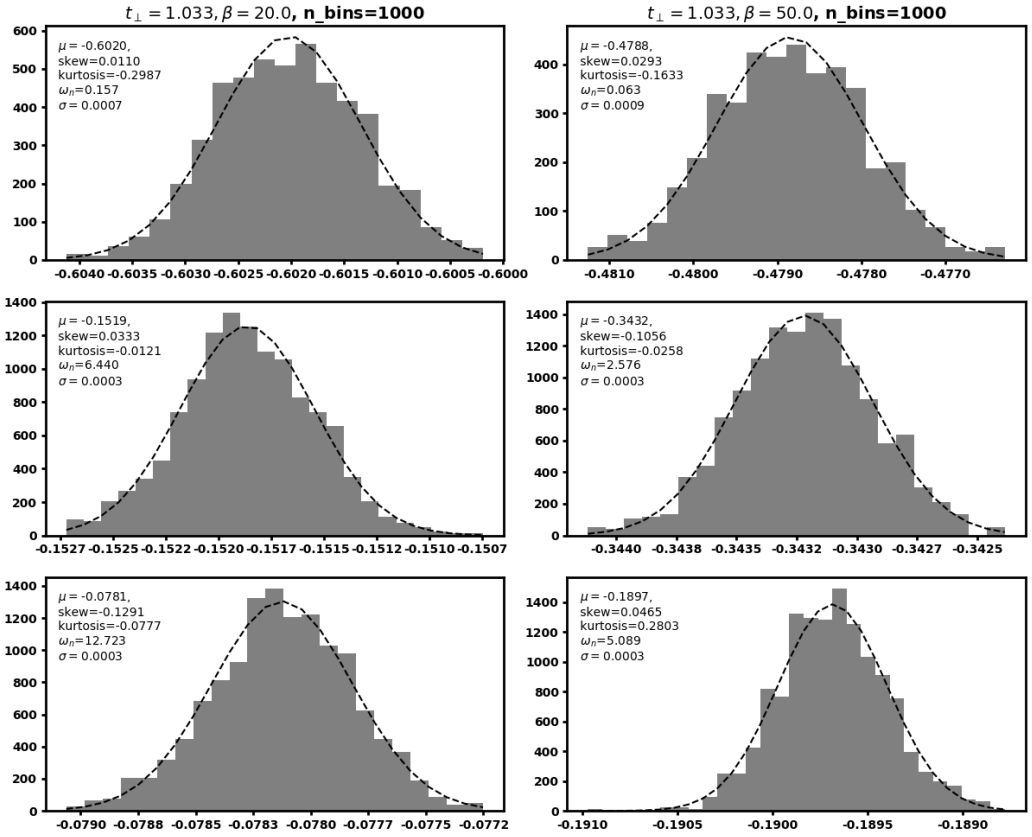


Figure 2: Histograms of the real part of the Green's function $G(i\omega_n)$ for four representative Matsubara frequencies, calculated from $n_{\text{bins}} = 1000$ independent CTQMC runs. The left panel is for $\beta = 20$ and the right for $\beta = 50$. The distributions are well-described by a Gaussian fit, confirming the statistical quality of the data. Other parameters are $t_\perp = 1.033, U = 1.75, V = 0.44$.

Figs. 2 and 3 show the distribution of points$(G(i\omega_n))$ for $n_{bins} = 1000$ and 2000, respectively, at different inverse temperatures$(\beta = 20, 50)$. We observe that the distribution appears Gaussian in both cases. The mean $\mu$ and standard deviation $\sigma$ of the distribution are nearly identical for both. In order to check the variation of skewness and kurtosis, we

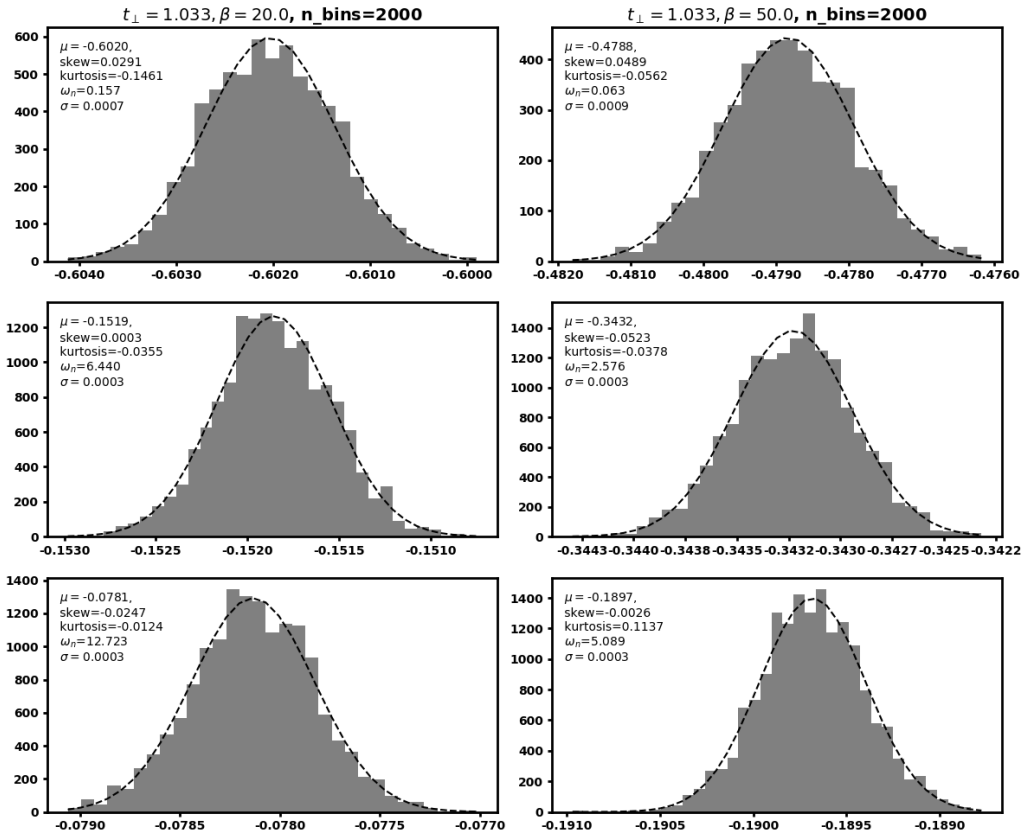


Figure 3: Histograms of the real part of the Green's function $G(i\omega_n)$ for four representative Matsubara frequencies, calculated from $n_{\text{bins}} = 2000$ independent CTQMC runs. The left panel is for $\beta = 20$ and the right for $\beta = 50$. The distributions are well-described by a Gaussian fit confirming the statistical quality of the data. Other parameters are $t_\perp = 1.033, U = 1.75, V = 0.44$.

plot them as a function of $n_{\text{bins}}$, as shown in Fig. 4. We observe that both the skewness and kurtosis have larger values for smaller sample sizes ($n_{\text{bins}}$) and decrease as the sample size increases. They reach reasonably low values when $n_{bins}$ 750. We conclude that the sample size of 750 or more is sufficient for the MaxEnt calculations.

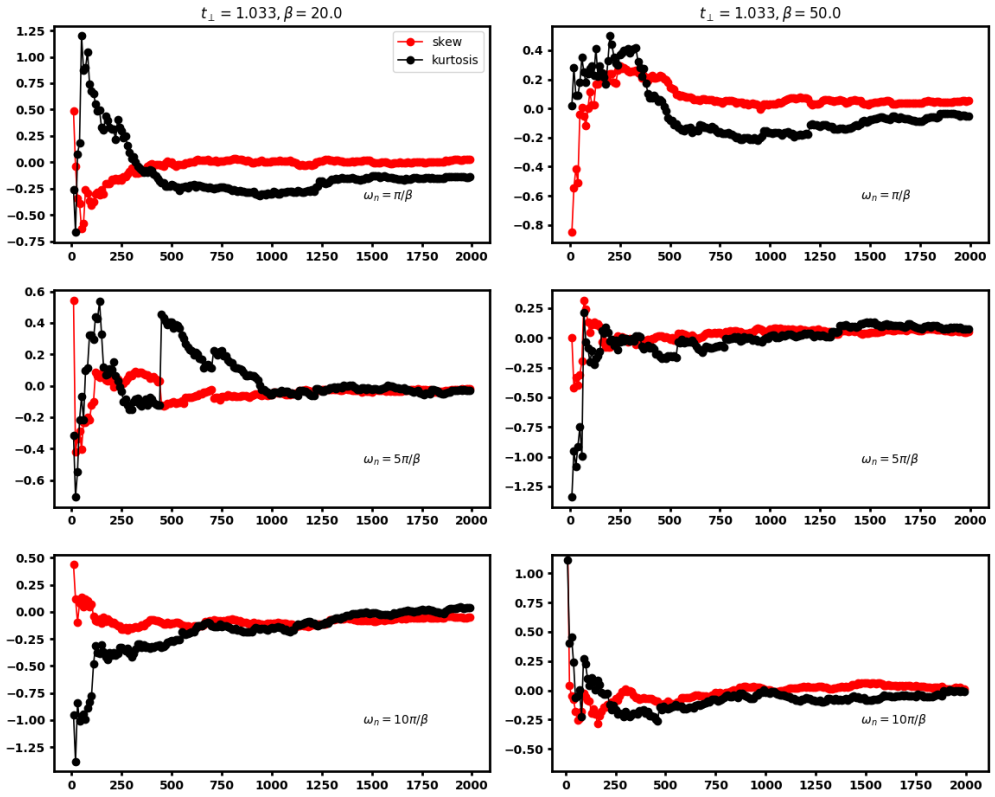


Figure 4: Skewness (red curve) and kurtosis (black curve) as a function of the number of statistical bins ($n_{\text{bins}}$) for three different Matsubara frequencies. Both moments approach zero for $n_{\text{bins}} \geq 750$, indicating that this sample size is sufficient to obtain normally distributed data for the MaxEnt analysis.

From the above analysis, we conclude that the data obtained from our calculations follow a normal distribution, confirming the absence of correlations between different datasets. However, there may still be correlations within each dataset. (Note that we carefully chose the length_cycle parameter to reduce such correlations [17].) To eliminate correlations among the data points within a dataset, we compute the covariance matrix and apply a unitary transformation to represent the Green's function in the diagonal basis of the covariance matrix.

In order to find out whether there is correlation within the data(means correlation between the same set of data but with different times), we construct a gradient plot of the covariance matrix as shown in the Fig. 5.

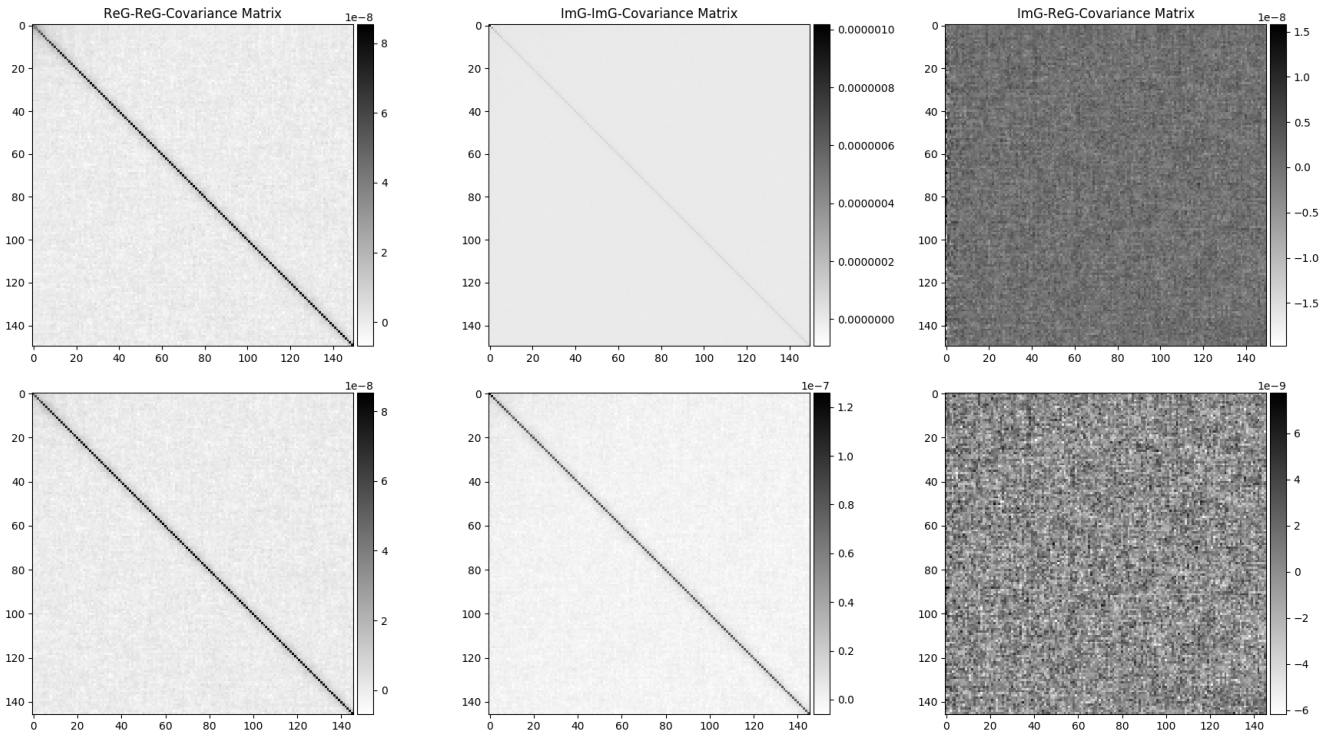


Figure 5: Fig. shows the covariance matrix of the Matsubara frequecy Green's function, where a) shows the covariance between Re$G(i\omega_n)$ and Re$G(i\omega_n)$, b) Im$G(i\omega_n)$ and Im$G(i\omega_n)$ and c) Im$G(i\omega_n)$ and Re$G(i\omega_n)$for the parameters $t_\perp = 1.033$, and $\beta = 20$.

## 1.4 Results

We determined the spectral function for various values of the interlayer coupling and $\beta$. Fig.6 shows how the spectral function changes with the change in $t_\perp$. Fig. 7 shows the variation of the spectral function with temperature for different interlayer coupling values. Note these calculations are done using triqs_maxent.

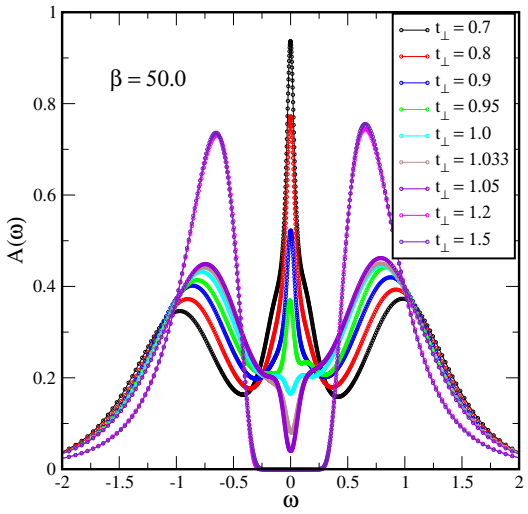


Figure 6: Plot of spectral function for $\beta = 50$ and for various $t_\perp$s.

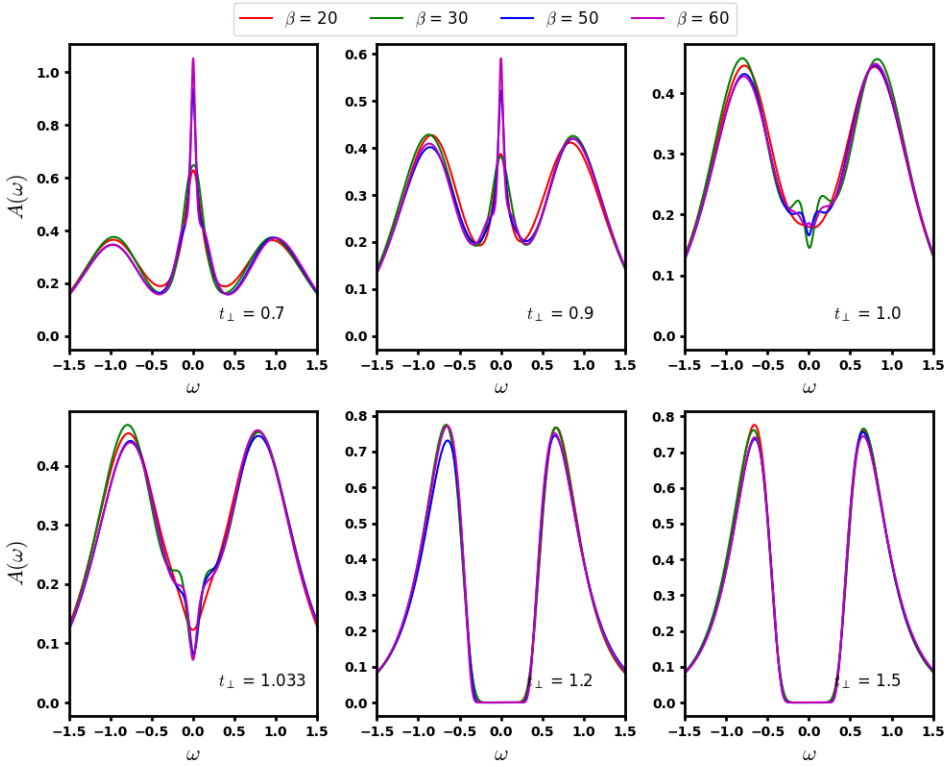


Figure 7: Plot of spectral function of the 3-orbital model. Different subplots are for different inter-layer couplings. $\beta$ is varied in each subplots.

### Analytic continuation of the $\Sigma(i\omega)$

Here we perform the analytic continuation for the self-energy using Omega_maxent. As shown in Figs. 8 and 9, the imaginary part of the self-energy becomes diverging as $t_\perp$ becomes higher, which is consistent with the paper (Phys. Rev. B 107, 205104 and arXiv:2509.00225).

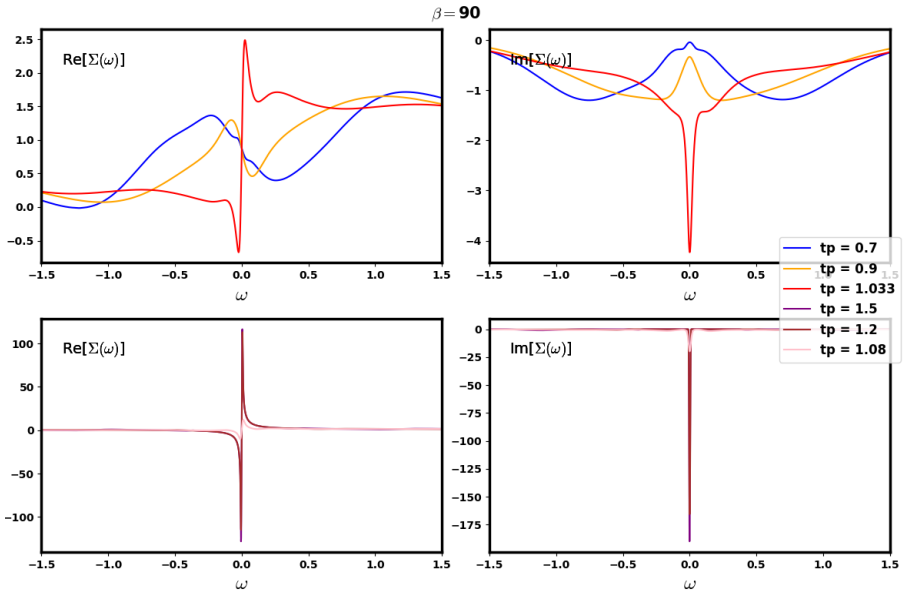


Figure 8: Self-energy calculated from CTQMC data using MaxEnt, is plotted for different $t_\perp$s and for $\beta = 90$

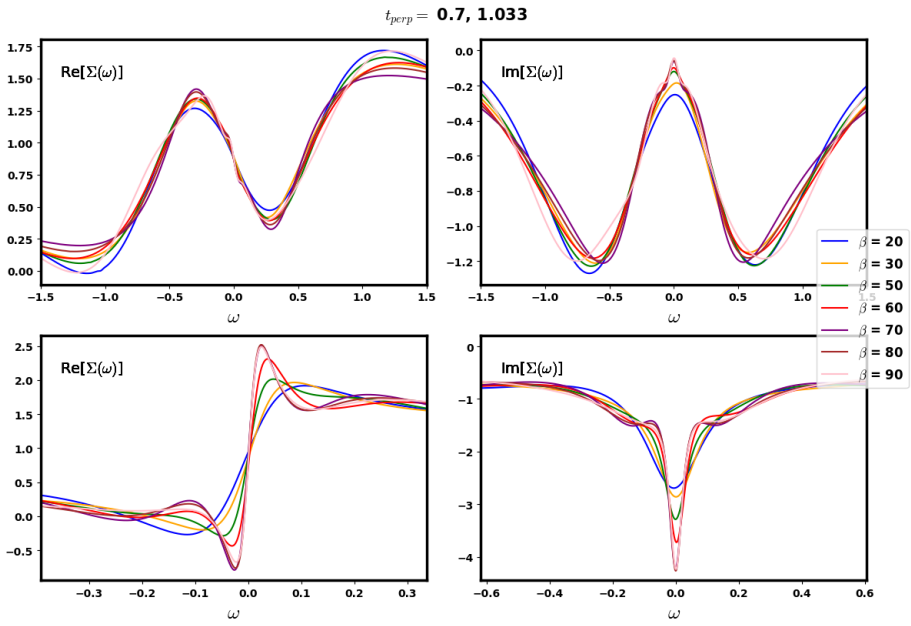


Figure 9: Self-energy calculated from CTQMC data using MaxEnt, is plotted for different $\beta$s and for $t_\perp = 0.7$(top-panel), $1.033$(bottom-panel)

# References

[1] H. Vidberg and J. Serene, Journal of Low Temperature Physics **29**, 179 (1977).

[2] J. Schött, I. L. Locht, E. Lundin, O. Grånäs, O. Eriksson, and I. Di Marco, Physical Review B **93**, 075104 (2016).

[3] J. Fei, C.-N. Yeh, and E. Gull, Physical Review Letters **126**, 056402 (2021).

[4] C. E. Creffield, E. Klepfish, E. R. Pike, and S. Sarkar, Physical review letters **75**, 517 (1995).

[5] M. Jarrell and J. E. Gubernatis, Physics Reports **269**, 133 (1996).

[6] D. Bergeron and A.-M. Tremblay, Physical Review E **94**, 023303 (2016).

[7] R. N. Silver, D. S. Sivia, and J. E. Gubernatis, Physical Review B **41**, 2380 (1990).

[8] R. Bryan, European Biophysics Journal **18**, 165 (1990).

[9] G. J. Kraberger, R. Triebl, M. Zingl, and M. Aichhorn, Physical Review B **96**, 155128 (2017).

[10] R. Fournier, L. Wang, O. V. Yazyev, and Q. Wu, Physical Review Letters **124**, 056401 (2020).

[11] R. Zhang, M. E. Merkel, S. Beck, and C. Ederer, Physical Review Research **4**, 043082 (2022).

[12] H. Yoon, J.-H. Sim, and M. J. Han, Physical Review B **98**, 245101 (2018).

[13] M. Kliczkowski, L. Keyes, S. Roy, T. Paiva, M. Randeria, N. Trivedi, and M. M. Maśka, Physical Review B **110**, 115119 (2024).

[14] E. T. Jaynes, Physical review **106**, 620 (1957).

[15] S. F. Gull and J. Skilling, in *Iee proceedings f (communications, radar and signal processing)*, Vol. 131 (IET, 1984) pp. 646–659.

[16] J. Skilling, in *Maximum Entropy and Bayesian Methods: Cambridge, England, 1988* (Springer, 1989) pp. 45–52.

[17] P. Seth, I. Krivenko, M. Ferrero, and O. Parcollet, Computer Physics Communications **200**, 274 (2016).